Contents - Index

HYGINIST for Windows Help Contents

Contents - Index

Getting Started (Summary)

A statistical tool like HYGINIST is useful if the effectiveness of working condition control measures is assessed using discontinuous exposure measurements.

HYGINIST includes 6 transactions:
1. data entry and mutation,
2. examining distribution shape,
3. estimating the descriptive statistics,
4. extrapolation to unsampled periods and compliance test against a limit value,
5. comparing the exposure data with other sets of descriptive statistics,
6. establishing the minimum sample size for an unbiased estimate of long-term compliance control.

This Help further explains:

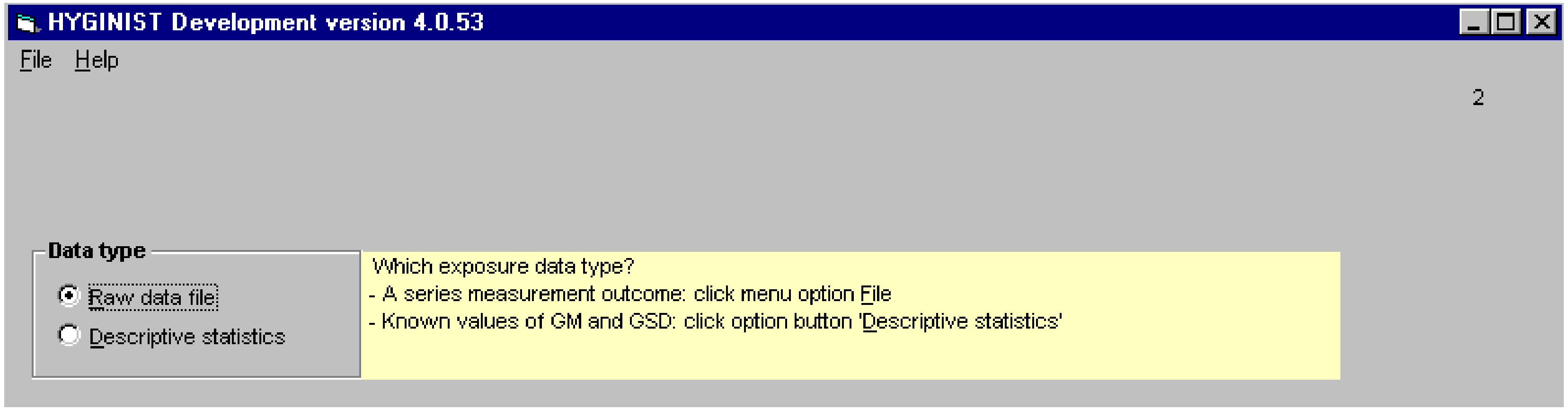
- How to apply the HYGINIST statistical methods in working conditions control
- The many aspects of the Lognormal distribution
- Installing and running the program
- Program operations
- Register for regular use.

.

If HYGINIST is started, then the following Splash screen appears.
After installation it will include the next Login box.

See Annex C.5.1. on how to fill out this form is

If the Login form is filled out correctly then the HYGINIST start page appears after a few seconds, in which all form and constants are loaded. Now you can start with the exposure data management and analysis.



I hope this program support your needs.
Please inform me by mail ihpc@planet.nl on your experience and don't forget checking regularly on the HYGINIST homepage for updates of the program and the helpfile.

Regards
Theo Scheffers

Click here for a summary on the Goodness-of-fit methods used.
Click here for a summary on the Extrapolation methods used.

HYGINIST for Windows Help Contents

- Figures

HYGINIST for Windows Help Index

Contents - Index

Accuracy

The difference between the measurement and the truth (Hawkins 1991), the extent in which calculations or estimations, the true value approximate

B.2.4.3 The linear estimator

The most simple, method to estimate the arithmetic mean of the Lognormal distribution uses Formula B-3 with $x_i=c_i$. This estimator of the mean is consistent, which means unbiased for infinite sample size.

Contents - Index

B.2.1 Mean and standard deviation

All estimators of EXP(μ) and EXP(σ) use the following formulas for mean and standard deviation:

$$\bar{x} = \frac{\sum_{i=1}^{M} x_i}{M}$$ B-3

$$s = \sqrt{\frac{\sum_{i=1}^{M} \left(x_i - \bar{x}\right)^2}{M-1}}$$ B-4

both with:

$$x_i = \log(c_i)$$ b-5

Contents - Index

B.2.4 Location (2). The arithmetic mean

The arithmetic mean β (first moment):

- has the units of measurement of concentration (or intensity),
- can take every real value over zero (by definition).

In the Lognormal distribution β is a function of EXP(μ) and EXP(σ):

$$\beta = EXP\left(\mu + \frac{\sigma^2}{2}\right) \quad \text{B-15}$$

The arithmetic mean β of a series:

- grab sample measurements within a reference period is a measure for the TWAreference period,
- more days TWA8 hour measurements is a measure for the mean dose of accumulating agents (Seixas 1988) like asbestos, silica, cadmium et cetera.
-

There are four sample estimators:

- The maximum likelihood (see B.2.4.5)
- The unbiased (see B.2.4.1)
- the minimal mean square error (see B.2.4.2)
- The linear (see B.2.4.3)

Confidence limits around the arithmetic mean estimator are calculated in B.2.4.4

B.2.4.5 The maximum likelihood estimator

Contents - Index

B.2.4.1 The unbiased estimator

The 'uniformly minimum variance unbiased' estimator (UMVU; Shimizu 1988b p 29, Finney 1941) of the arithmetic mean AM of the Lognormal distribution can be derived, in the same way as was done for GM_u, in a Taylor series expansion (see Shimizu 1988b p30 & 31 formula 2.3. & 2.6.):

$$AM = GM * \phi\left(\frac{s^2}{2}\right)$$

with

$$\Phi(t) = 1 + \frac{(M-1)*t}{M} + \frac{(M-1)^3*t^2}{M^2*(M+1)*2!} + \frac{(M-1)^5*t^3}{M^3*(M+1)*(M+3)*3!} + ... \quad \text{B-16}$$

Although well known in industrial hygiene (Oldham 1953, Galbas 1975, Bar-Shalom 1975 part II page 47; 1976 page 472, Leidel 1977, Owen 1980) its use is inhibited because of the believed possible deterrent to its users (Dewell 1993) and/or misinterpretation (Selvin 1989).

The superior characteristics of formula 5.16 over other estimators (e.g. the maximum likelihood and direct estimators) are described by Kuo-Hsing Chang (1990) and Attfield (1992, 1993). formula B-16 converges much more quickly to a constant value than formula B-11. For s^2>.1 the relative increase in the 7^e or 8^e term is, for all values of MD2, less than 1E-8. In BASIC, formula B-16 is programmable in a small subroutine.

Contents - Index

B.2.4.2 The minimal mean square error estimate

The estimator with the smallest mean square error among the estimators (Zhou 1998) is:

$$AM = GM * \phi\left(\frac{(M-3) * s^2}{2 * M}\right)$$

with

$$\Phi(t) = 1 + \frac{(M-1) * t}{M} + \frac{(M-1)^3 * t^2}{M^2 * (M+1) * 2!} + \frac{(M-1)^5 * t^3}{M^3 * (M+1) * (M+3) * 3!} + \ldots \quad \text{B-28}$$

This estimator can only be applied for sample sizes of at least four.

B.2.4.4 Confidence limits

Exact confidence limits exist (Land 1971, Bar-Shalom 1975). They can be established (Armstrong 1992, see Example 33) using the tables (Land 1975) or can be calculated as such (Land 1988 6.1). HYGINIST calculates the exact limits but also the approximate method that was used in HYGINIST version 2.2 and earlier.

See example 51

Figure 23 Rankit plot of 2000 AM values. Figure 24 Rankit plot of 2000 AM values. Monte Carlo Lognormal. Sample size 50 Monte Carlo Lognormal. Sample size 50

Contents - Index

5.2.1 Uniform most powerful

The 'uniform most powerful' (=UMP) test for the arithmetic mean of a Lognormal distribution is developed by Land (1971, 1988 p89 & 103), introduced in the industrial hygiene by Galbas (1975) and Bar-Shalom (1975, 1976), and supported by Leidel (1977 page 55) and Coenen (1978). The UMP is:

- most highly preferred because it is unbiased with minimum variance,
- difficult in its use, because it can only be solved analytically, for sample size odd.

HYGINIST for Windows contains the algorithm to calculate the UMP estimator of:
- the probability that the mean exceeds the limit
- the Upper confidence limit of the mean

if:
- Sample size is 3<M<1001
- confidence 90%<D<99.75%

Example 32 The UMP test for the arithmetic mean of a Lognormal distribution

Contents - Index

Example 32 The UMP test for the arithmetic mean of a Lognormal distribution

*.HYG file Description

HAW117 With 14 random $TWA_{8\ hour}$ PAS total dust above LL=1.4 mg/m^3 the chronic health hazard was tested. The plotting positions $y={}_{10}log(GM/H)=-0.66$ and $s={}_{10}log(GSD)=0.08$ (see screen 24a) falls inside compliance area of Nomogram 4.3 in Leidel (1977 page 58). Since GSD is in part explained by CV_t (see Example 57), the situation is in real compliance.

LEIDEL56 With 8 $TWA_{20\ min}$ Ethyl alcohol PAS data the $TWA_{8\ hour}$ is classified against H=1000 PPM (Leidel 1977 page 56). Entering y=0.002 and s=0.14 in the nomograms, results that the situation is classified as possible overexposure (like Leidel concluded 1977 p.57). Because the location parameter GM is independent of CV_t, overexposure is real even without good knowledge of the corrected GSD (see Example 57).

BAR_SI25 In an industrial plant 6 grab sample measurements of carbon monoxide were performed for a particular employee at the hours 9, 10, 11, 12, 14, and 16. (Bar-Shalom 1975 page I-25). If the resulting y= -0.02 and s=0.22 are entered in the nomograms, then the situation is classified as possible overexposure against a TWA_{8hour}=50 PPM.

OWEN716 The 15 grab sample airborne Chlorine concentrations are used to classify for a $TWA_{8\ hour}$ standard of 1 PPM. After introducing a two-sided accuracy range of 0.25-9 PPM the rankit estimators are GMg=.88 PPM, GSDg=6.81 and df=10. Extrapolation towards the plotting positions y=-0.00 and s=0.67 for df+1=11 in Leidel (1977 figure 4.3.). indicates possible over exposure (Owen 1980 p716)

Contents - Index

Table 1 Example data from different sources

*.HYG file agent Sample Limit dimension Source
size M value H
ALBRE220 TWA8 hr Aspartame 4 ? µG/M3 Albrecht (1989) p220
ACN8_9 TWA8 hr Acrylonitril 116 4 PPM Swaen (1992)
BAR_SI25 Grab Sample Carbon monoxide 6 50 PPM Bar-Shalom (1975) pI-25
BILAN304 TWA8 hr Methylene bisphenyl 5 0.005 PPM Bilan (1989) p304 sprayer indoor
isocyanate (MDI)
BOLEY62 TWA8 hr Total dust 12 10 MG/M3 Boleij (1987) p62, paper plant
BOLEY85 TWA8 hr ZnCl2 total Dust 7 mg/m3 Boleij (1987) p85
CHIP123 Lead 10 50 µG/M3 Booher (1988) p123
COHEN132 max. Susquehanna River flood levels 20 - 106*ft3/sec Cohen (1988) p132
CONOV195 two-digit telephone numbers 50 - - Conover (1980) p365
DEWELL24 TWA8 hr Respirable dust 8 5 MG/M3 Dewell (1989) p24
DEWELL42 TWA10 min Formaldehyde 10 2C PPM Dewell (1989) p42
DEWELL44 Welding fumes 11 5 MG/M3 Dewell (1989) p44
GUPTA271 Mice survival time 10 - days Gupta (1952) p271 &
Schneider (1986) p69 & 88
HALD151 Diameters of rivet heads 500 - mm Hald (1952) p151
HAW104 TWA8 hr Chlorine 10 0.5 PPM Hawkins (1991) 104
HAW117 TWA8 hr Total dust 15 10 MG/M3 Hawkins (1991) 117
LEIDEL56 TWA20 min Ethyl alcohol 8 1000 PPM Leidel (1977) p56-60
LEIDEL61 Grab sampled Ozone 35 0.1 PPM Leidel (1977) p61-62
LEIDEL63 TWA10 min Hydrogen Sulphide 5 20C PPM Leidel (1977) p63-64
LEIDEL67 TWA8 hr Dioxane 10 100 PPM Leidel (1977) p67-69
LEIDL103 Grab sampled HF 12 3C PPM Leidel (1977) p103-4
LD103_10 HF minus background 0.1 ppm 12 3C PPM Leidel (1977) p103-4
LEIDL104 TWA8 hr Methylmethacrylate 24 100 PPM Leidel (1977) p104
MOF134NS TWA10-30 min Methoxyfluorane 16 - PPM Potts (1988) p134 Nonscavenged
MOF134S TWA10-30 min Methoxyfluorane 6 - PPM Potts (1988) p134 Scavenged
OWEN716 Grab sampled Chlorine 15 1 PPM Owen (1980) p 716
POSTB11I TWA8 hr Styrene 14 100 PPM Post (1989) A11 Inlayers 08 thru 14
POST1008 TWA8 hr Styrene 9 100 PPM Inlayer/press 10/4 and 08/15
POST08P4 TWA8 hr Styrene 5 100 PPM Inlayer 08 press 4
POST10P5 TWA8 hr Styrene 4 100 PPM Inlayer 10 press 5
RANDOM20 Monte Carlo 2000 - - Lognormal distribution
EXP(µ)=1, EXP(σ)=2.71828
SARH212 90Sr in milk 10 - nCu/M3 Sarhan (1962) p212
SCHNE224 Items under stress 50 1 time units Schneider (1986) p224
SCHNE70 Failure distance of locomotives 96 ? 103 miles Schneider (1986) p70
SOLV198 TWA8 hr Hydrocarbons 45 575? MG/M3 Scheffers (1985) p19
X07-10 Chemical X 30 - PPB Gustafson (1991) ch7

Contents - Index

F.1 Computer program and quantitative evaluation

Contents - Index

F.1.1 English references

Abramowitz M., Stegun I.A. Handbook of mathematical functions. New York: Dover Publications Inc. (1964) 5e edition 1970.
ACGIH Protection of the sensitive individual; papers of a symposium, Cicinnati, Nov 1981. Ann. Am. Conf. Ind. Hyg. 3 (1982).
ACGIH 1993-1994 Treshold Limit Values for chemical substances and physical agents and Biological Exposure Indices Cincinnatti 1993.
d`Agostino R.B. An omnibus test of Normality for moderate and large size samples. Biometrika 58 (1971) 341-348.
d`Agostino R.B. Small sample probability points for the D test of Normality. Biometrika 59 (1972) 219-221.
Armstrong B.G. Confidence intervals for the arithmetic means of Lognormally distributed exposures. Am. Ind. Hyg. Assoc. J. 53 (1992) 481-485.
Attfield M.D., Hewett P. Exact expressions for the bias and the variance of estimators of the mean of a Lognormal distribution. Am. Ind. Hyg. Assoc. J. 53 (1992) 432-435.
Attfield M.D., Hewett P. Letter to the editor. Am. Ind. Hyg. Assoc. J. 54 (1993) A-86.
Bar-Shalom Y., et al Handbook of statistical tests for evaluating exposure to air contaminants. HEW Publication No. (NIOSH) 75-147 Washington: U.S. Department of Health, Education and Welfare (1975).
Bar-Shalom Y., Segall A., Budenaers D. Decision and estimation procedures for air contaminants. Am. Ind. Hyg. Assoc. J. 37 (1976) 469-473.
Beasley J.D., Springer S.G. The percentage points of the Normal distribution. Applied Statistics 26 (1977) 118-121.
Berry G., Day, N.E. The statistical analysis of the results of sampling an environment for contaminant when most samples contain an undetectable level. Am. J. of Epidemiology 97 (1973) 160-166.
Boneau C.A. The effects of violations of assumptions underlying the t test. Psychological Bulletin 57 (1960) 49-64.
Brief R.S., Scale R.A. Occupational exposure limits for novel work schedules. Am. Ind. Hyg. Assoc. J. 36 (1975) 467-469.
Brief R.S., Scala R.A. Occupational Health Aspects of Unusual Work Schedules: A Review of Exxon Experiences. Am. Ind. Hyg. Assoc. J. 47 (1986) 199-202.
Buring E., Lanting R. Working paper on monitoring strategies in the United Kingdom, Federal Republic of Germany and United States of America. Labour Inspectory S 69. Voorburg 1989.
Buringh E., Lanting R. Exposure variability in the workplace: Its implications for the assessment of compliance. Am. Ind. Hyg. Assoc. J. 52 (1991) 6-13.
CEFIC Report on occupational exposure limits and monitoring strategy. Conseil Européen des Federations de l`industrie Chemique. Brussel (no publication year mentioned, probably 1984).

CEN/T 137 Workplace atmospheres - Guidance for the assessment of exposure to chemical agents for comparison with limit values and measurement strategy. EN 689. Brussels (1995).
CEN/T 137 Workplace atmospheres - General requirements for the performance of procedures for measurement of chemical agents Draft prEN 482. Brussels (1995).
Checkoway H., Pearce N., Crawford-Brown D.J. Research methods in occupational epidemiology. Monographs in epidemiology and biostatistics. Oxford university press 13 (1989).
Cicollela A., Servais M., Vincent R., Wild P. STARTREX (STAtistique, Resultats, TRaitement EXploitation) Logiciel de traitement des données d'exposition aux substances chimiques. Developpé par: Institut National de Recherche et de Sécurité (1993).
Clayton G.D., Clayton F.E. Patty's industrial hygiene and toxicology. Third revised edition. Volume I General principals. John Wiley & Sons (1978).
CONCAWE Review of strategies for the evaluation of employee exposures to substances hazardous to health. CONCAWE's Industrial Hygiene Subgroup. Report no. 87/57 The Hague October 1987.
Corn M. Strategies of air sampling. Scand. j. Work Environm. Health 11 (1985) 173-180.
Crow E.L. Minimum variance unbiased estimators of the ratio of means of two Lognormal variates and of two gamma variates. Commun. Statist.-Theor. Meth. A6 (1977) 967-975.
Damiano, J. A guideline for managing the industrial hygiene sampling function. Am. Ind. Hyg. Assoc. J. 50 (1989) 366-371.
Dewell P. Some applications of statistics in occupational hygiene. BOSH Technical Handbook Series No. 1 (1989).
Dewell P. Letter to the editor. Am. Ind. Hyg. Assoc. J. 54 (1993) A85-A86.
EG The council of the european communities. Council directive of 16 december 1988 (88/642/EEC). Official Journal of the European Communities No L 356/74. 24.12.88.
Eide I. The application of 8-hour occupational exposure limits to non-standard work schedules off-shore, Ann. Occup. Hyg. 34 (1990) 13-17.
Eller P.M. Manual of analytical methods Third edition U.S. Department of Health and Human services NIOSH Cincinnati (1984 with updates).
Esmen N.A, Hammad Y.Y. Log-normality of environmental sampling data. J. Environ. Sci. Health. 162 (1977) 29-41.
Esmen N.A. A distribution-free double-sampling method for exposure assessment. Appl. Occup. Environ. Hyg. 7 (1992) 613-621.
Evans J.S., Hawkins N.C. The distribution of Student's t-statistics for small samples from Lognormal exposure distributions Am. Ind. Hyg. Assoc. J. 49 (1988) 512-515.
Filiben J.J. The probability plot correlation coefficient test for Normality. Technometrics 17 (1975) 111-117.
Finney D.J. On the distribution of a variate whose logarithm is Normally distributed. J. Royal Statistical Society Suppl., 7 (1941) 155-161.
Fisher R.A., Yeates F. Statistical tables for biological, agricultural and medical research. New York: Hafner (1938) 6e edition 1963 Table 20.
Francis M., et al The effect of autocorrelation on the estimation of workers' daily exposure. Am. Ind. Hyg. Assoc. J. 50 (1989) 37-43.
Gaddum, J.H. Lognormal distributions. Nature 156 (1945) 463-366.
Gerlach B. A correlation-type goodness-of-fit test for normality with censored sampling. Math. Operationsforsch. Statist., Ser. Statistics 11 (1980) 207-218.

Gupta A.K. Estimation of the mean and the standard deviation of a Normal population from a censored sample. Biometrika 39 (1952) 260-273.
Geary R.C. Testing for Normality. Biometrika 34 (1947) 209-242.
Gustafson T.L. TRUE EPISTAT Manual. Fourth edition. Epistat services. Texas (1991).
Guest G.I., Cherrie J.W., Gardner R.J., Money C.D. Sampling strategies for airborne contaminants in the workplace. British Ocuupational Hygiene Society Technical committee working group. Technical Guide no. 11 (1993)
Hald A. Statistical Theory with engineering applications. par.. 6.9.: The truncated Normal Distribution. Wiley (1952) 144-151.
Hall I.J., Sampson C.B. One-sided tolerance limits for a Normal population based on censored samples. J. Statist. Comput. Simul. 2 (1973) p 317-324.
Harter H.L. Expected values of Normal order statistics. Biometrika 48 (1961) 151-165.
Harter H.L. Order statistics and their use in testing and estimation. Vol 2: Estimates based on order statistics of samples from various populations. Chapter III: Expected values of order statistics of samples from various populations. par.. 1: Normal population. Washington U.S. Government Printing Office (1970) 25-29.
Harris R.L. Guideline for collection of industrial hygiene exposure assessment data for epidemiologic use. Chemical Manufacturers Association (1993) 36p.
Hawkins N.C. Evans J.S. Subjective estimation of toluene exposures: A calibration study of industrial hygienists. Applied Industrial Hygiene 4 (1989) 61-68.
Hawkins N.C., Norwood S.K., Rock J.C. (ed.) A strategy for occupational exposure assessment. AIHA Akron, Ohio (1991) 179p.
Hawkins N.C., Jayjock M.A., Lynch J. A rationale and a framework for establishing the quality of human exposure assessments. Am. Ind. Hyg. Assoc. J. 53 (1992) 34-41.
Heeren T., d'Agostino R. Robustness of the two independent samples t-test when applied to ordinal scaled data. Statistics in medicine 6 (1987) 79-90.
Hickey J.L., Reist P.C. Application of occupational exposure limits to unusual Work Schedules. A. Ind. Hyg. Assoc. J. 38 (1977) 613-621.
Hill I.D. The Normal integral. Applied Statistics 22 (1973) 424-427.
Hornung R.W., Reed L.D. Estimation of average concentration in the presence of Nondetectable Values. Appl. Occup. Environ. Hyg. 5 (1990) 46-51.
HSE Monitoring strategies for toxic substances. Health & Safety Executive. Guidance note EH 42. Environmental Hygiene 42 (1989) 1-12.
HSE Methods for the Determination of Hazardous Substances. MDHS ... Health and Safety Executive. i.e.: Methods for the Determination of Hazardous Substances. MDHS 67: Total (and speciated) chromium in chromium plating mists. Health and Safety Executive, february 1990
ISO International Standard: Workplace Air - Determination of ...., ISO,...., Genève. i.e.: International Standard: Workplace Air - Determination of particulate lead and lead compounds - Flame atomic absorption spectrometric method. ISO 8518, december 1990.
Jahr J. Calculations for workplace measurements and other data with a Lognormal distribution. Staub. Reinhaltung der Luft. 47 (1987) 153-156.
Johnson N.L., Kotz S. Distributions in statistics. Continuous univariate distributions-1 Chapters 12, 13 & 14 (1970) 1-136.
Johnson N.L., Kotz S. Distributions in statistics. Continuous univariate distributions-2 Chapters 26, 27, 30 & 31 (1970) 75-129, 189-219.

Kendall M.G. The advanced theory of statistics Volume II (1947).
Kendall M.G., Buckland W.R. A dictionary of statistical terms. 4th edition revised and enlarged. International Statistical Institute 1982.
Kerr G.W. Use of statistical methodology in environmental monitoring. Am. Ind. Hyg. Assoc. J. 23 (1962) 75-82.
Kromhout H., et al Agreement between qualitative exposure estimate and quantitative exposure measurements. Am. J. Ind. Medicine 12 (1987) 551-562.
Kromhout H., Symanski E., Rappaport A.M. A comprehensive evaluation of within- and between- worker components of occupational exposure to chemical agents. Ann. occupat. Hyg. 17 (1993) 253-270.
Kuo-Hsing Chang, Nelson P.I. Letter to the editor. Am. Ind. Hyg Assoc. J. 51 (1990) A-813.
Land C.E. Confidence intervals for linear functions of the Normal mean and variance. Ann. Math. Statistics 42 (1971) 1187-1205.
Land C.E., Johnson B.R., Johsi V.M. A note on two-sided confidence intervals for linear functions of the Normal mean and variance. The Annals of Statistics 1 (1973) 940-943.
Land C.E. Tables of confidence limits for linear functions of the Normal mean and variance. In: *Selected tables in mathematical statistics. Harter H.L. and Owen D.B. eds.*, Volume III. American Mathematical Society (1975) 385-419.
Land C. E. Hypothesis tests and interval estimation. In: *STATISTICS: Textbooks and Monographs Volume 88 Lognormal Distributions. Edited by Edwin L. Crow and Kunio Shimizu.* New York: Marcel Dekker (1988) 87-112.
Laurent A.G. The Lognormal distribution and the translation method: Description and estimation problems. Am. Stat. Assoc. J. 58 (1963) 231-235.
Leidel N.A., Bush K.A. Statistical Methods for the determination of non-compliance with Occupational Health Standards (1975a). HEW Publication No. (NIOSH) 75-159 Cincinnati: U.S. Department of Health, Education and Welfare.
Leidel N.A., Bush K.A., Crouse W. E. Exposure measurement action level and occupational environmental variability (1975b). HEW Publication No. (NIOSH) 76-131 Cincinnati: U.S. Department of Health, Education and Welfare.
Leidel N.A., Bush K.A., Lynch J.R. Occupational exposure sampling strategy manual. DHEW, NIOSH publ. 137, (1977) Cincinnati: U.S. Department of Health, Education and Welfare.
Mietinnen O.S. Theoretical epidemiology: principles of occurrence research. New York: John Wiley & Sons (1985)
Mehran F. Variance of the MVUE for the Lognormal mean. JASA 68 (1973) 726-727.
Morrison J. The Lognormal distribution in quality control. Applied Statistics 7 (1958) 160-172.
Mulhausen J.R., Damiano J. A strategy for Assessing and managing Occupational Exposures. AIHA Fairfax, Viginia (1998) 349p.
Nicas M., Simmons B.P., Spear R.C. Environmental versus analytical variability in exposure measurements. Am. Ind. Hyg. Assoc. J. 52 (1991) 553-557.
NIOSH Manual of Analytical Methods, third edition. U.S. Department of Health and Human Services.
Odeh R.E., Owen D.B. Tables for Normal tolerance limits, sampling plans, and screening. Statistics: textbooks and monographs. Vol 32 (1980) 17-69.
Oldham P.D. The nature of the variability of dust concentration at the coal face. Brit. J. Ind.

Med. 10 (1953) 227-234.
Oldham P.D. On estimating the arithmetic means of Lognormally-distributed populations. Biometrics 39 (1965) 235-239.
OSHA Analytical methods manual. Vol. VI - Industrial hygiene technical manual. OSHA Analytical Laboratory, Salt Lake City, Utah, 1985
Owen D.B. Tables for computing bivariate Normal probabilities. Ann. of Mathematical Statistics 27 (1956) 1075-1095.
Owen D.B. A special case of the bivariate Noncentral t-distribution. Biometrika 52 (1965) 437-446.
Owen D.B. A survey of properties and applications of the Noncentral t-distribution. Technometrics 10 (1968) 445-477.
Owen W.J., DeRouen T.A. Estimation of the mean for Lognormal data containing zeroes and left-censored values, with applications to the measurement of worker exposure to air contaminants. Biometrics 36 (1980) 707-719.
Page E.B. Ordered hypotheses for multiple treatments: a significance test for linear ranks. American Statistical Association Journal (1963) 216-230
Patel J.K., Read C.B. Handbook of the Normal distribution. Chapter 7: Normal approximations to distributions. Statistics: textbooks and monographs. Vol 40 (1982) 168-223.
Patel J.K., Read C.B. Handbook of the Normal distribution. Chapter 10: The bivariate Normal distribution. Statistics: textbooks and monographs. Vol 40 (1982) 288-329.
Pearson E.S., d'Agostino R.B., Bowman L.R. Test for departure from Normality: Comparison of powers. Biometrika 64 (1977) 231-246.
Perkins J.L., Cutter G.N., Cleveland M.S. Estimating the mean, variance an confidence limits from censored (<limit of detection), Lognormally-distributed exposure data. Am. Ind. Hyg. Assoc. J. 51 (1990) 416-419.
Post W., et al Qualitative estimates of exposure to methylene chloride and styrene: the influence of quantitative exposure data. Applied Industrial and Environmental Hygiene 6 (1991) 179-204.
Prescott P. Estimation of the standard deviation of a Normal population from double censored samples. Biometrika 57 (1970) 409-419.
Press W.H., et al Numerical recipes. The art of scientific computing. Cambridge (1987) p229 en p724.
Proschan F. Confidence and tolerance intervals for the Normal distribution. Am. Stat. Assoc. J. 48 (1953) 550-564.
Rappaport S.M. Assessment of long-term exposure to toxic substances in air. Ann. Occup. Hyg. 35 (1991) 61-121.
Rappaport S.M., Selvin S. A method for evaluating the mean exposure from a Lognormal distribution. Am. Ind Hyg. Assoc. J. 48 (1987) 374-379.
Rappaport S.M., Spear R.C. Physiological damping of exposure variability during brief periods. Ann. Occup. Hyg. 32 (1988) 21-33.
Rappaport S.M., Kromhout H., Symanski E. Variation of exposure between workers in homogeneous exposure groups. Am. Ind. Hyg. Assoc. J. 54 (1993) 654-662.
Rempel D. Occupational Medicine. Medical surveillance in the workplace. Philadelphia (1990) 435.
Roach S.A. A most rational basis for air sampling programmes Ann. Occup. Hyg. 20 (1977)

65-84.
Roach S.A., et al Testing compliance with threshold limit values for respirable dusts. Am. Ind. Hyg. Assoc. J. 28 (1967) 543-553.
Rock J.C. A comparison between OSHA-compliance criteria and action-level decision criteria. Am. Ind. Hyg. Assoc. J. 43 (1982) 297-313.
Royston J.P. An extension of Shapiro and Wilk's W test for Normality to large samples. Applied Statistics 31 (1982a) 115-124.
Royston J.P. Expected Normal order statistics (exact and approximate). Applied Statistics 31 (1982b) 161-165.
Royston J.P. The W-test for Normality. Applied Statistics 31 (1982c) 176-180.
Royston J.P. Correcting the Shapiro-wilk W for ties. J. Stat. Comput. Simul. 31 (1989) 237-249.
Royston P. Estimating departure from Normality. Statistics in Medicine 10 (1991) 1283-1293.
Royston P. Estimation, reference ranges and goodness of fit for the three-parameter log-normal distribution. Statistics in Medicine 11 (1992a) 897-912.
Royston P. Approximating the Shapiro-Wilk W-test for non-Normality. Statistics and Computing 2 (1992b) 117-119.
Royston P. A toolkit for testing for non-normality in complete and censored samples. The Statistician 42 (1993) 37-43.
Royston J.P. Remark AS R94. A remark on Algorithm AS 181: The W-test for Normality. Applied Statistics 44 (1995) 547-551.
Sarhan A.E., Greenberg B.G. Estimation of location and scale parameters by order statistics from single and double censored samples, Part 1, The Normal distribution up to samples of size 10. Ann. Math. Statist. 27 (1956) 427-451.
Sarhan A.E., Greenberg B.G. Contributions to Order Statistics. New York: John Wiley (1962). Chapter 10. Normal Distributions.
Scheffers T.M.L., Jongeneelen F.J., Bragt P.C. Development of effect-specific limit values (ESLVs) for solvent mixtures in painting. Ann. Occup. Hyg. 29 (1985) 191-199.
Schneider H. Truncated and censored samples from Normal populations. Statistics: textbooks and monographs. Vol 70 (1986).
Schneider H. Simple and highly efficient estimators for censored Normal samples. Biometrika (1984) 412-414.
Selvin S., Rappaport S.M. A note on the estimation of the mean value from a Lognormal distribution. Am. Ind Hyg. Assoc. J. 50 (1989) 627-630.
Seixas N.A., Robins T.G., Moulton L.W. The use of the geometric and arithmetic mean exposures in occupational epidemiology. Am. J. Ind Medicine 14 (1988) 465-477.
Shapiro S.S., Wilk M.B. An analysis of variance test for Normality. Biometrika 52 (1965) 591-611.
Shapiro S.S., Wilk M.B., Chen H.J. A comparative study of various tests for Normality. Am. Stat. Assoc. J. 63 (1968) 1343-1372.
Shapiro S.S., Francia R.S. An approximate analysis of variance test for Normality. J. Am. Stat. Assoc. 67 (1972) 215-216.
Shapiro S.S., Brain C.W. A review of distributional testing procedures and development of a censored sample distributional test. In *Statistical distributions in scientific work* Editors C. Taillie et. al. Reidel Publshing Compagny 5 (1981) 1-24.

Shapiro S.S. How to test Normality and other distributional assumptions. -2nd ed. Volume 3 in *The ASQC basic references in quality control: Statistical techniques* Editors S.S. Shapiro and E.F. Mykytka. The American Society for Quality Control (1990).
Shimizu K., Crow E.L. History, Genesis, and Properties. In: *STATISTICS: Textbooks and Monographs Volume 88 Lognormal Distributions. Edited by Edwin L. Crow and Kunio Shimizu.* New York: Marcel Dekker (1988a) 1-25.
Shimizu K. Point Estimation. In: *STATISTICS: Textbooks and Monographs Volume 88 Lognormal Distributions. Edited by Edwin L. Crow and Kunio Shimizu.* New York: Marcel Dekker (1988b) 27-86.
Siegel S. Nonparametric statistics for the behavioral sciences. McGraw-Hill New York (1956).
Smith T.J. Exposure assessment for occupational epidemiology. Am. J. Ind Medicine 12 (1987) 249-268.
Snedecor G.W., Cochran W.G. Statistical methods. The Iowa State University press. (1980) 7e edition.
Spear R.C., Selvin S., Francis M. The influence of averaging time on the distribution of exposures. Am. Ind. Hyg. Assoc. J. 47 (1980) 365-386.
SPSS SPSS/PC+? Update for V3.0 and V3.1 for the IBM PC/XT/AT and PS/2. Chicago (1989).
Stellingwerf J. Time dependent aspects in the european communities directives of asbestos and lead. Letter to the editor. Br. J. Ind. Med. 41 (1984) 354-356.
Tait K. The workplace exposure assessment expert system (WORKSPERT). Am. Ind. Hyg. Assoc. J 53 (1992) 84-98.
Teichroew D. Tables of expected values of order statistics and products of order statistics for samples of size twenty and less from the Normal distribution. Annals of Mathematical Statistics 27 (1956) 410-425.
Thomas A.A. Skin Absorption: a potential contributing factor to the BEI (biological exposure index). Appl. Ind. Hyg. 1 (1986) 87-90.
Travis C.C., Land M.L. Estimating the mean of data sets with nondetectable values. Environ. Sci Technol. 24 (1990) 961-962.
Tuggle R.M. The NIOSH decision scheme. Am. Ind. Hyg. Assoc. J 42 (1981) 493-498.
Tuggle R.M. Assessment of occupational exposure using one-sided tolerance limits. Am. Ind. Hyg. Assoc. J 43 (1982) 338-346.
Verrill S., Johnson R.A. The asymptotic equivalence of some modified Shapiro-Wilk statistics - complete and censored sample cases. The Annals of Statistics 15 (1987) 413-419.
Waters M.A. A measure of goodness-of-fit for the Lognormal model applied to occupational exposures. Am. Ind. Hyg. Assoc. J. 52 (1991) 493-502.
WHO Health effects of combined exposures in the work environment. Report of a WHO Expert Committee. Technical Report Series 662. World Health Organization Geneva (1981).
WHO Evaluation of exposure to airborne particles in the work environment. WHO offset publication no 80. World Health Organization Geneva (1984).
Wilks S.S. Determination of sample sizes for setting tolerance limits. Ann. Math. Statistics 12 (1941) 91-96.

F.1.2 Non English references

Bezemer P.D. Referentiewaarden. Een verkenning voor het bepalen van `normale waarden`. Proefschrift VU 1981.
BIN Strategie voor het meten van stoffen met een tijdgewogen grenswaarde. Belgisch Instituut voor Normalisatie NBN T96-002. juni 1987.
Boleij J., Heederik D., Kromhout H. Karakterisering van blootstelling aan chemische stoffen in de werkomgeving. Wageningen: Pudoc 1987.
Buringh E., Huygen C. Concept voorlichtingsblad meetstrategie 4. Frequentieverdeling van meetgegevens. Toetsing en schatting. TNO-MT Rapport R 88/251b Delft 1989.
Ciba-Geigy Wissenschaftliche Tabellen Geigy. Basel (1980).
Coenen W. Zur Frage des Vertrauensbereiches bei Mittelwerten der Staubkonzentration. Staub Reinhalt. Luft 26 (1966) 216-221.
Coenen W., Riediger G. Die Schatzung des zeitlichen Konzentrationsmittelwertes gefährlicher Arbeidsstoffe in de Luft bei stichprobenartigen Messungen. Staub- Reinhaltung der Luft 38 (1978) 402-409.
DFG Analytischen Methoden zur Prüfung gesundheitsschadlicher Arbeitsstoffe. Band 1/3, Luftanalysen. Deutsche Forschungsgemeinschaft.
DFG idem, Band 2/3, Analysen in biologischen Material.
Galbas H.G. Schätz- und Testmethoden für den Erwartungswert einer Lognormal verteilten Grundgesamtheit. Staub- Reinhaltung der Luft 35 (1975) 448-452.
Heederik D., et al Meetstrategieën voor chemische belasting ten behoeve van risico-analyze. NVvA symposium verslag. Den Haag: Nederlandse Vereniging voor Arbeidshygiëne (1987).
Jonge H. de Inleiding tot de medische statistiek. Deel 1. Fundamentele begrippen en technieken verdelingsvrije methoden. Nederlands Instituut voor Praeventieve Geneeskunde, Leiden (1963).
Jonge H. de Inleiding tot de medische statistiek. Deel 2. Klassieke methoden. Nederlands Instituut voor Praeventieve Geneeskunde, Leiden (1964).
Juda J., Budzinski K., Bestimmung des Toleranzbereiches des Mittelwertes bei den Messungen der Staubkomzentration. Staub Reinhaltung der Luft 27 (1967) 176-179.
Manz A., Manz J. Praktische Beispiele für die Toxicologische Bewertung der Gesamtwirkung verschiedener gesundheitsgefahrdender Stoffe in der Luft am Arbeitsplatz. Zentralblatt für Arbeitsmedizin, Arbeitsschutz, Prophylaxe und Eergonomie 37 (1987) 278-284.
NNI Luchtkwaliteit - Werkplek atmosphere. Bepaling van de concentratie aan ...., NVN, ... Delft. Bijvoorbeeld: Luchtkwaliteit - Werkplek atmosphere. Bepaling van de concentratie aan arseen in stof door middel van atomaire-absorptiespectrometrie na actieve monsterneming op een membraanfilter. Nederlandse Voornorm, NVN 2951, juni 1990.
Oostendorp Y., Kromhout H. Protokol voor algemeen chemisch werkplekonderzoek. Landbouw Hogeschool Wageningen Vakgroep, Luchthygiëne en -verontreiniging.

Wageningen juli 1985a.
Oostendorp Y., Kromhout H. Model voor werkplekonderzoek naar de blootstelling aan chemische stoffen. Landbouw Hogeschool Wageningen, Vakgroep Luchthygiëne en -verontreiniging. R-154. Wageningen November 1985b.
Scheffers T. EPO: een arbeidshygiënisch en gezondheidskundig informatie systeem van werknemers met een beroepsmatige blootstelling aan chemische stoffen. In: *Meetstrategieën voor chemische belasting ten behoeve van risico-analyze. NVvA symposium verslag. Redactie D.H. Heederik.* Nederlandse Vereniging voor Arbeidshygiëne, Den Haag (1987). 77-89.
Scheffers T. HYGINIST, handleiding bij het computerprogramma. NVvA Den Haag 1990, 100 blz. (inclusief computerprogramma).
Scheffers T., et al DOHS-Base. Een databestand met arbeidshygiënische grenswaarden en meetmethoden voor de Nederlandse arbeidshygiënist. Een publikatie van de werkgroep Grenswaarden en Meetmethoden van de Nederlandse Vereniging voor Arbeidshygiëne (Eerste uitgave, januari 1992) 1-43.
Stresemann E. Die allergologische Bedeutung von Stoffen der MAK-Werte-Liste. Zentralblatt für Arbeitsmedizin, Arbeitsschutz, Prophylaxe und Eergonomie 38 (1988) 389-393.
TRgA 402 Messung und Beurteilung von Konzentrationen gefährlicher Arbeitsstoffe in der Luft; Anwendung von Maximalen Arbeidsplatzkonzentrationen (MAK). Bundesarbeitsblatt 11 (1984) 55-61.
VDI 2450 Messen von Emission, Transmission und Immision luftverunreinigender stoffe. Methoden zur behandlung einzelner Variabeln. Quantile. Verein Deutscher Ingenieure (1979).

5.2.2 Approximation

Screen 24b displays the:
- probability that the arithmetic mean β is above limit H for dfD25 and GSDD1.64,
- confidence limits including the arithmetic mean β with confidence U%.

The approximate test of β against limit H is based on the following statistic (Jahr 1987, Armstrong 1992):

$$t_{df} = \frac{LOG\left(\frac{H}{AM}\right) * \sqrt{df}}{LOG(GSD)}$$

Formula 5-12

Deviate t_{df} is, under the null hypothesis, supposed to follow the Student distribution. Estimator $A\beta_{>H}$ of the right side chance $\alpha\beta_{>H}$ is calculated using t_{df}, df and the Student distribution. The corresponding upper confidence limit $Ç\beta_{,U}$ is derived from by logarithmic transformation:

$$C_{\beta,U\%} = AM * GSD^{\frac{t_{\alpha,df}}{\sqrt{df}}}$$

Formula 5-13

Example 33 Comparing different estimators of the confidence interval of the arithmetic mean

Contents - Index

Example 33 Comparing different estimators of the confidence interval of the arithmetic mean

Based on the maximum likelihood estimator AMML=exp(+s²/2)=1058 PPM, Armstrong (1992 table II) calculated the two-sided 95% confidence limits for the 8 $TWA_{20\ min}$'s Ethyl alcohol (LEIDEL56.HYG) using the exact (Land 1971) and four approximate methods. The bottom line is calculated with 5.13.

Method two-sided 95%-tile in ppm
Exact 824 1437
simple Student 800 1294
Lognormal Student 808 1386
Cox 841 13322
modified Cox 802 1397
5.13 788 1402
Because the sample has a normal shape and GSD is small and completely explained by CVt (see Example 57), it is difficult to draw any solid conclusion. The approximate method 5.12 seems to have a high sensitivity of finding possible overexposure but a low specificity of rejecting compliance.

Using one-sided values of confidence 2.5% and 97.5% The exact Land method provides in HYGINIST a confidence range of 831.4 thru 1470 ppm

Maximum likelihood

One of the methods that provides estimators of the parameters of a parent population from samples. ML estimators fulfil certain criteria for consistency, efficiency and sufficiency. In fact ML estimators are consistent, tend to Normality for large sample size, have minimum variance in the limit at least, and provide sufficient statistics where such exist (Kendall 1947 volume II par.. 17.22)

Contents - Index

Example 57 CVt Influence on GSD in example data

*.HYG file | Description

LEIDEL63 5 short period exposures to hydrogen sulphide provides an uncorrected relative standard deviation w/AM=0.11. Adjusting GSD=1.12 is not possible because w/AM < CV_t=0.12 (Leidel 1977 page 79 table D-1).

HAW117 After removing two results (because the local exhaust system had misfunctioned during sampling) and entering a lower detection limit at LL=1.4 mg/m^3 (outlier correction), the remaining 14 $TWA_{8\ hour}$ total dust (Hawkins 1991 p117) provides GSD_g=1.19 (Figure 11) and w/AM=0.18. Adjusting GSD_g for random measurement errors using df=13 and total dust CV_t=0.15, provides a GSD_{adj}=1.1

BOLEY62 Boleij (1987 page 63) corrected GSD for CV_t, in a series $TWA_{8\ hour}$ total dust from a papermill. That is why the GSD=2.68, calculated on the complete sample, differs from the GSD=2.5 presented by the author. Using the raw data, 9 out of 12 ex ceed a lower detection limit of 1.8 mg/m^3 (see Figure 12) providing GM_g=4.7 mg/m^3, GSD_g=3.35 and w/AM=1.33, which overshadows any reasonable total dust CV_t. So adjusting GSD in this sample is not necessary.

For LEIDEL67.HYG and HAW117.HYG the untransformed Shapiro probability A(W) gives the highest conformity of shape (see Example 15), a second indication that random error determined the variance.

Contents - Index

Example 15 Other transformations

Conformity with other distributions than the Lognormal is found in the following examples from Table 1:

*.HYG file Description
BILAN304 Bilan (1989 p304) indoor sprayer with MDI exposure: $A(W)_{log}$=43.7%, $A(W)_{lglg}$=83%,
DEWELL24 Dewell (1989 p24) dust in foundry: $A(W)_{log}$=34.3%, $A(W)_{lglg}$=82%,
LEIDEL56 Leidel (1977 p56) 8 $TWA_{20\ min}$ Ethyl Alcohol (Figure 10): $A(W)_{log}$=41.8%, $A(W)_{unt}$= 94.9%, $A(W)_{X^2}$=99.2%

Contents - Index

Example 51 Arithmetic mean

In Figure 23 and Figure 24 the cumulative distributions of 2000 AMs are displayed on Log-probability scale. AMs are calculated from a series of M=50 or M=2 Monte Carlo Lognormal deviates (μ=0, σ=1). Lognormal fit of the AM distribution improves with increasing sample size. The Lognormal approximation of the AM distribution can be described by the following descriptive statistics

- EXP(μAM) is nearly EXP($\mu+\sigma^2/2$)=1.6
- EXP(σAM) is nearly EXP(σ)+SQR(1/df) =2.72 (df=1) & 1.15 (df=49).

The rankits estimators displayed in Figure 22 and Figure 23 are comparable:

- GMgdf=1=1.7 and GSDgdf=1=2.73
- GMgdf=49=1.7 and GSDgdf=49=1.19

The 50th and 1950th values of the sample data of Figure 22 are 1.19 and 2.29, respectively. The 95% two-sided confidence interval of AM is:

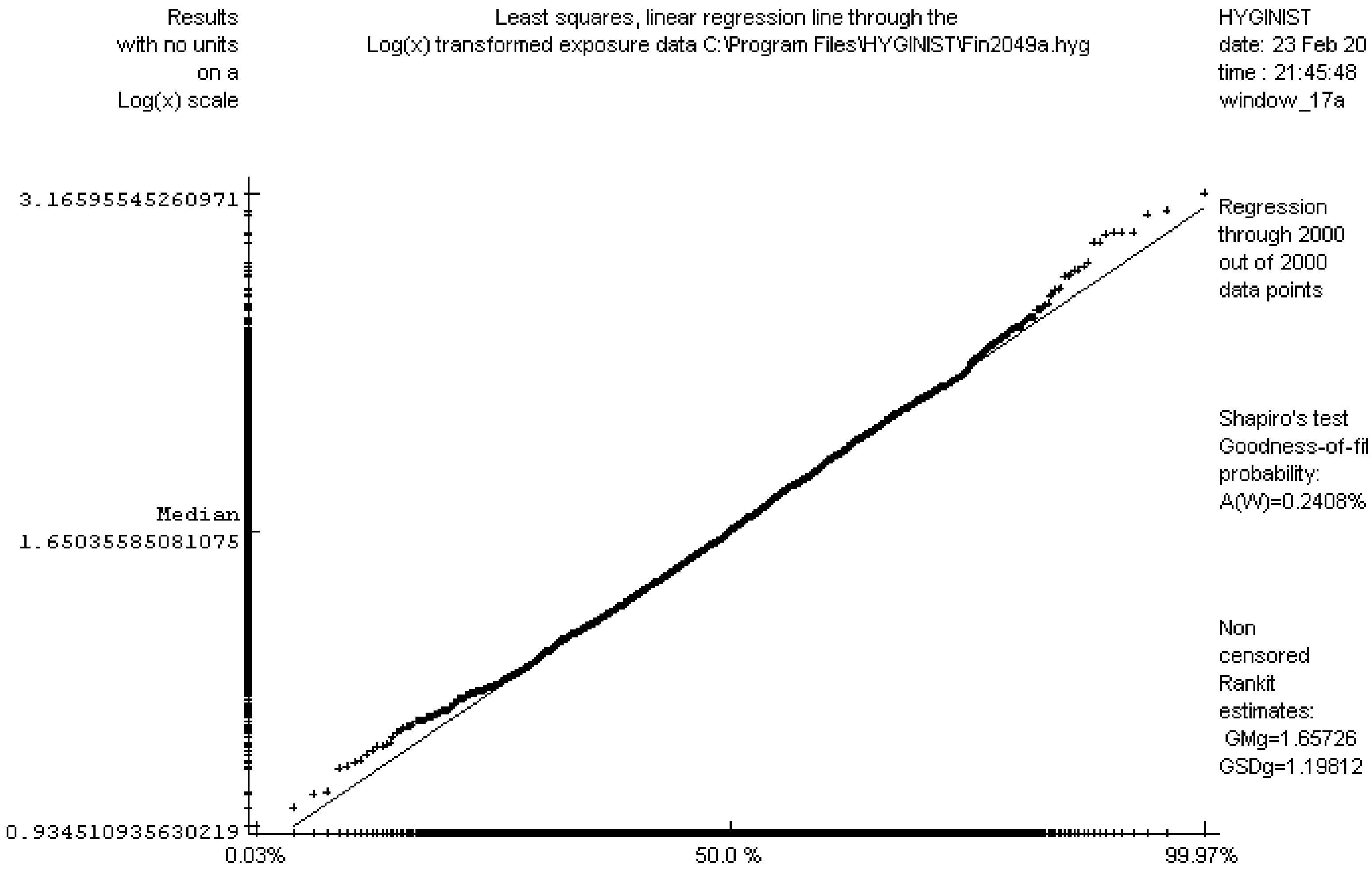
- $e^{0.5-2.149/\%(49)}$=1.21 and $e^{0.5+2.827/\%(49)}$=2.47, based on Land (1975 page 413; s=1, df=49), .
- 1.22 and 2.17, based on the approximate method of Jahr (1987) .

Figure 23 Rankit plot of 2000 AM values. Monte Carlo Lognormal. Sample size 50

Figure 24 Rankit plot of 2000 AM values. Monte Carlo Lognormal. Sample size 50

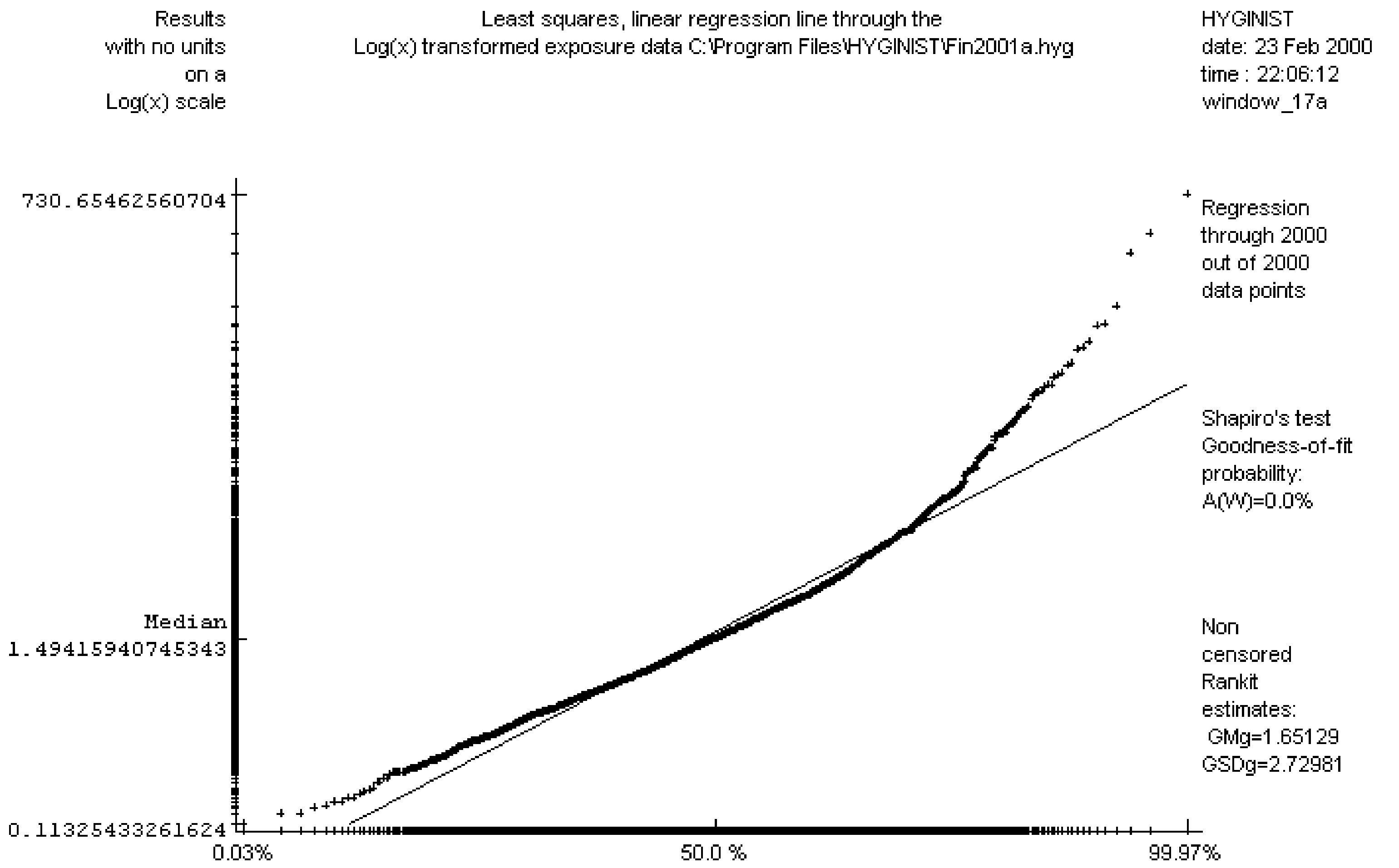
Contents - Index

Figure 23 Rankit plot of 2000 AM values. Monte Carlo Lognormal. Sample size 50



Contents - Index

Figure 24 Rankit plot of 2000 AM values. Monte Carlo Lognormal. Sample size 50



Contents - Index

B.2.5 Variance (2). The relative, arithmetic standard deviation

The variance (the second moment) of the Lognormal distribution is (Finney 1941 formula 4, Mehran 1973, Shimizu 1988b p9 formula 4.3 and Attfield 1992):

$$\omega^2 = \mathrm{EXP}(2\mu + \sigma^2) * [\mathrm{EXP}(\sigma^2) - 1] \quad \text{B-17}$$

The unbiased estimator with minimum variance is (Finney 1941 formula 15, Shimizu 1988b p31 formula 2.7):

$$w = GM * \left[\phi\left(\frac{s^2}{2}\right) - \phi\left(\frac{M-2}{M-1} * s^2\right) \right]^{1/2} \quad \text{B-18}$$

with $\phi(t)$ as formula B-16. The relative, arithmetic standard deviation (Shimizu 1988 page 10) is:

$$\frac{\omega}{\beta} = \sqrt{\exp\left(\sigma^2\right) - 1} \quad \text{B-19}$$

formula B-19 is independent of the location and is also called the sample coefficient of variation (Shimizu 1988b formula 4.10). Its estimator w/AM is the ratio of formula B-18 and formula B-16. w and w/AM are used in the comparison with the:

- measurement coefficient of variation CV_t,
- the grouping interval -C.

If CV_t and -C/AM are of the same order of magnitude as w/AM, then the Lognormal shape of the exposure variability will be camouflaged.

The most simple method to estimate the standard error of the Lognormal distribution uses the formula B-4 with $x_i = c_i$.

Contents - Index

Asymptotic unbiased

Asymptotic unbiased or consistent estimator: An estimator which expectation equals the parameter by infinite sample size

Autocorrelation

the phenomenon that observations in a series are dependent upon the preceding

Contents - Index

Averaging time

a period of time for which the measuring procedure yields a single result (prEN 482)

Contents - Index

Bias

systematic error (Hawkins 1991 156). Consistent deviation of the results from the true value (ISO 6879)

Contents - Index

Censored distribution

A distribution with a known fraction of observations outside the detection range (Hald 1949)

B.3.1 Censored sample estimators for EXP(m)

In the case of a censored sample the program estimates μ from the linear regression through the uncensored data M' in the rankits plot. GM_g is the regression line median, corresponding to the 50%-tile on the x-axis (R_j=0):

$$GM^g = EXP\left(\frac{\sum_{i=ll}^{i=ul}\left[x_i - LOG(GSD^g)*R_i\right]}{M'}\right)$$

In formula formula B-21 *ll* and *ul* are the lowest and highest uncensored data rank and *M'* is the number of results between the accuracy limits. Formula B.21 is identical with the linear, alternative estimators of Gupta (1952 5.2. formulas 31 through 35) and Sarhan (1962 p208 formulas 10C.2.1 and 10C.2.2). See Example 14, 21 and 44 on the rankit estimators for mean and variance of strontium in milk.

Example 52 Comparing untransformed linear estimators

For uncensored samples the rankit estimator is as effective as the classical estimators (see B.2). They are more effective than classical solution where results below the detection lower limit LL receive the value .5*LL (Hornung 1990, Hawkins 1991 p 104). Example 53 suggest that the rankits estimators should be preferred in industrial hygiene.

Contents - Index

Example 14 Comparing xcg and scg with estimators from literature

*.HYG file Description

HALD151 The rankit estimators of the untransformed diameters of rivet heads (=13,4304 mm
and s=0,1098 mm Figure 24) are in close agreement with the ML estimators
(=13,429 mm and s=0,111 mm) of Hald (1952). Both estimations are
based on the M`=310 rivet heads above the lower detection limit of 13.40 mm.
The number of degrees of freedom is df=310+190/2-1=404.
The standard deviation s_{cg}=0.11 mm overshadows the grouping interval -C=0.05 mm.
Because of the small relative standard deviation s_{cg}/x_{cg}=.00812, the Lognormal
model may be effective as well, with GM_g=13.4 mm and GSD_g=1.00815.

SARH212 Students measuring Strontium-90 concentrations in milk.
Introducing both a lower limit of LL=8 pCu/l and an upper limit of
10 pCu/l results in M`=5 uncensored data out of M=10 (see Figure 25).
The untransformed rankit estimators of μ and σ are x_{cg}=9.33 pCu/l and s_{cg}=1.87 pCu/l
which are equal to the values calculated by Sarhan (1962 page 212).

HAW104 10 TWA's chlorine (.HYG). 3 Below LL=.05 PPM. The untransformed Normal probability
plot (Figure 8) shows an almost perfect regression through the uncensored data.
The Normal descriptive statistics are x_{cg}=.1948 PPM and s_{cg}=.1992 PPM.

Contents - Index

Example 21 Comparing GMg and GSDg with estimators from literature

*.HYG file Description

SCHNE224 GMg=9.1 and GSDg=2.12 (see Figure 5) of items under stress are in close agreement with the biased corrected, restricted ML estimators GM=$10^{.942}$=8.7 TU and GSD=$10^{.3123}$=2.05 (Schneider 1988 p 224).

LEIDL104 Leidel (1977 p105) estimates GM=34 PPM and GSD= 1.9 from the Lognormal probability plot on 24 TWA8 hour Methyl Methacrylate concentrations in the job category "Mix man". Introducing a virtual lower limit of LL=8 PPM in Screen 17d (see Figure 21), results in GMg=34.5 PPM and GSDg=1.91 PPM based on the complete M`=24 data.

Contents - Index

Example 52 Comparing untransformed linear estimators

Students measuring Strontium-90 concentrations in milk with µ=9.22 picoCuries per litre. The 2 lower and the 3 upper extremes (outliers) were excluded from the 10 observations (SARH212.HYG). From the untransformed, double censored sample the rankit regression estimators are mean =9.3251 and standard deviation s=1.8679. These values are equal to the alternative estimators =9.33 and s =1.87 and comparable with the exact estimators x=9.29 and s=1.69 (Sarhan 1962).

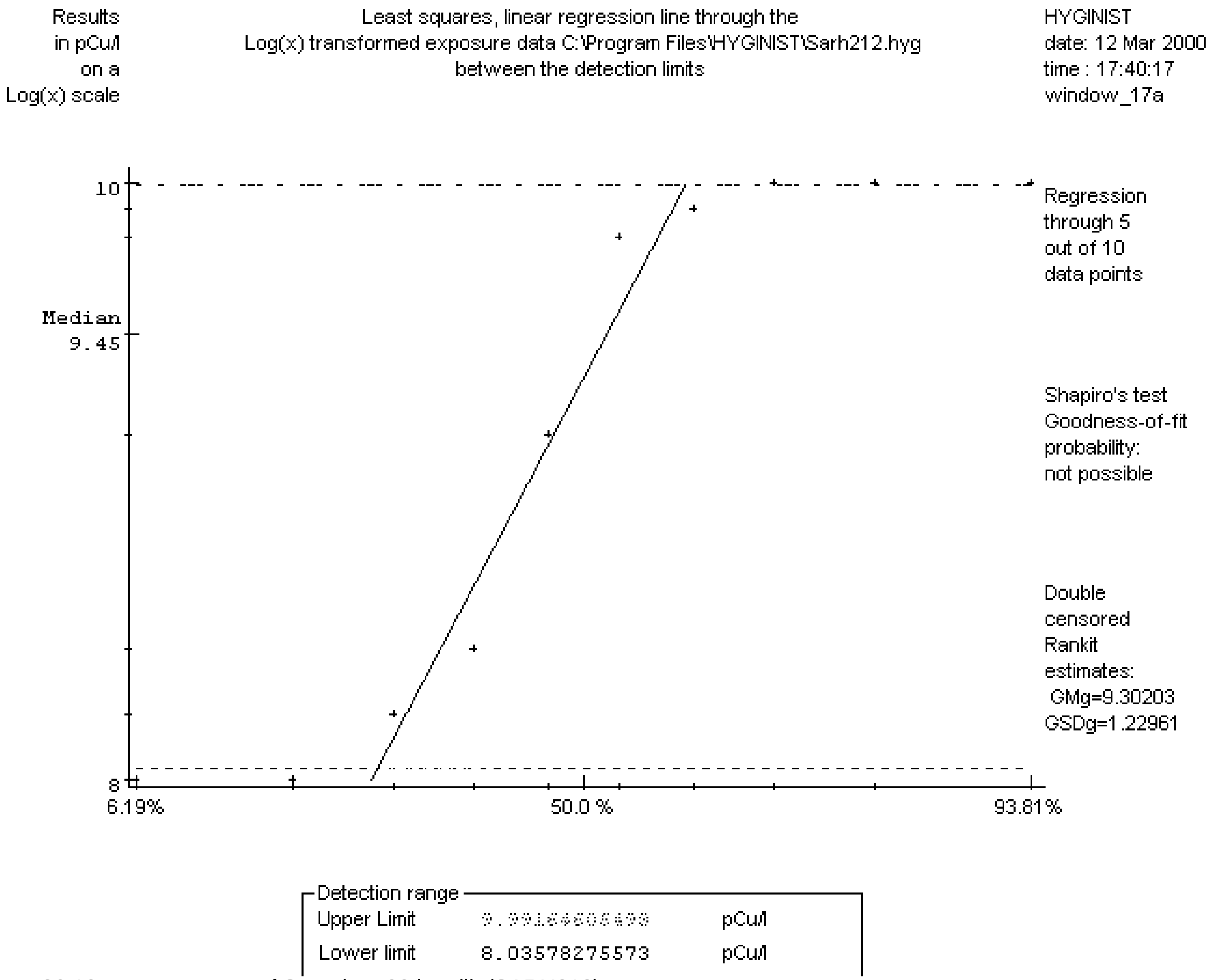


Figure 26 10 measurements of Strontium-90 in milk (SARH212)

Contents - Index

B.2 Parametric descriptive statistics (complete sample)

The Lognormal probability density function is completely described by its descriptive statistics μ and σ2 (Shimizu 1988 page 2):

$$g(C) = \frac{1}{\sigma * C * \sqrt{2\pi}} \mathrm{EXP}\left(- \frac{1}{2} * \frac{[\mathrm{LOG}(C) - \mu]^2}{\sigma^2} \right) \quad \text{for} \quad C > 0$$

$$g(C) = 0 \quad \text{for } C \leq 0 \qquad \text{B-2}$$

Table 17 displays the symbols used for the population parameters and sample estimators, and refers to the formulas and pages where the algorithms are displayed.

Descriptive statistics are measures for:

- location (modus, median, geometric and arithmetic mean, fractions, extreme values) (B.2.2 & B.2.4),
- variance (arithmetic and geometric standard deviation, range) (B.2.3 & B2.5), and
- shape (degrees of freedom, GSD, the higher moments of the normal distribution like kurtosis and skewness).

Descriptive statistics are normally divided in two groups:

- parametric descriptive statistics, which can describe completely a continuous probability density function: the (Log)normal mean and standard deviation s, the Weibull variance and shape, the Chi-square shape df (Shapiro 1991 page 6 & 8),
- distribution free descriptive statistics, which are independent of the shape and thus characterize the probability density function only in part: the extremes, median, modus, fractions and the number of results above or below the industrial hygiene limit value.

The relation between the four measures of location of the Lognormal distribution is: Modus < Median = Geometric mean Exp(μ) < Arithmetic mean β.

See example 50

Contents - Index

Table 17 Notations used for the Lognormal descriptive statistics

Name Descriptive statistic formula & page
Population parameter Sample estimates uncensored censored
Measures for location
Mean of log(C) μ formula B-3 -

Geometric Mean EXP(μ) GM, GMu, GMg formula B-9 & formula B-21 & formula B-11 formula B-11

Arithmetic Mean β=EXP($\mu+\sigma^2/2$) AM formula B-16 formula B-16

Median EXP(μ) ME - -

Modus EXP($\mu-\sigma^2$) - - -

Measures for scale
Variance of log(C) σ^2 s^2 formula B-4 -

Geom. Standard Dev. EXP(σ) GSD, GSDg formula B-13 formula B-20

Arithmetic Standard- ω=EXP($\mu+\sigma^2/2$)* w formula B-18 formula B-18
Deviation (ASD) sqr{EXP(σ^2)-1}

Relative ASD ω/β=sqr{EXP(σ^2)-1} w/AM formula B-19 formula B-19

Contents - Index

Modus

The value in a distribution with the highest frequency.

Contents - Index

Median

The most central result or, in case that sample size is even, the arithmetic mean of the untransformed two most central data

B.2.2 Location (1). The geometric mean

The geometric mean EXP(μ):

- is a measure of source strength (Seixas 1988, Rock 1982),
- has a unit dimension of intensity or concentration,
- can take every real value over zero,
- is identical to the median.

If an estimator of EXP(μ) exceeds a corresponding hygienic limit value H, then at least one result of the sample exceeds limit H.

There are four sample estimators:

- the maximum likelihood (see B.2.2.3)
- The unbiased (see B.2.2.1)
- The median
- The rankit estimator

For the confidence limits around GM see B.2.2.2

Notes:

- GM in B.6 is biased but consistent: for limited sample size it overestimates EXP(μ) on the average. See B.2.2.1.
- If the sample include undetectables then GM is estimated using rankit regression B.3.1.
- GM is necessary for the calculation of:
- tests on the limit value (chapter 5), and
- ~~inference statistics (chapter 6).~~
- sample size calculations (chapter 7)

B.2.2.3 The maximum likelihood estimator

The most commonly used estimator for EXP(μ) for a sample without undetectables is the antilog of the sum of the logarithms of the results c_i, divided by the sample size M:

$$GM = \mathrm{EXP}(\bar{x}) \quad \text{B-6}$$

with $\bar{x}$ defined as in formula B.3.

Contents - Index

B.2.2.1 The unbiased estimator

The bias of the Geometric Mean estimator GM (see B.6) is exactly known (Laurent 1963 formula 1):

$$E(GM) = EXP\left(\mu + \frac{\sigma^2}{2M}\right) \quad \text{B-7}$$

The average overestimation EXP(σ^2/2M) decreases with increasing sample size M. For σ^2 about unity, a common value for the variance, and M<=10 the bias is >=5%. An unbiased estimator of EXP(μ) with minimum variance was developed by Laurent (1963, see Shimizu 1988 page 29). Laurent's algorithm (1963, formula 7), however, is a Bessel function of the first kind and order (M-3)/2, and is difficult to calculate with values for the descriptive statistics which are relevant in industrial hygiene.

The following algorithm is more effective in industrial hygiene. The bias function EXP(-σ^2/2M) is first expanded in a Taylor series (Abramowitz 1970 formula 4.2.1.):

$$EXP\left(\frac{-\sigma^2}{2M}\right) = \sum_{k=0}^{\infty} \frac{\left(-\frac{\sigma^2}{2M}\right)^k}{k!}$$

$$= 1 - \left(\frac{\sigma^2}{2M}\right) + \left(\frac{\sigma^4}{(2M)^2 * 2!}\right) - \left(\frac{\sigma^6}{(2M)^3 * 3!}\right) + \ldots \quad \text{B-8}$$

In the Taylor series the unbiased and efficient estimator of the numerator σ^{2k} is (Laurent 1963 formula 4):

$$E(\sigma^{2k}) = \frac{\Gamma\left(\frac{M-1}{2}\right)}{\Gamma\left(k + \frac{M-1}{2}\right)} * s^{2k} * \left(\frac{M}{2}\right)^k \quad \text{B-9}$$

The ratio of the two Gamma functions G[(M-1)/2] and G[k+(M-1)/2] can be expanded and easily programmed in a series using formula 6.1.22 of Abramowitz (1970):

$$\frac{\Gamma_M}{\Gamma_{k,M}} = \frac{1}{\left(k - 1 + \frac{M-1}{2}\right) * \left(k - 2 + \frac{M-1}{2}\right) * \ldots * \left(\frac{M-1}{2}\right)} \qquad (k > 0) \quad \text{B-10}$$

Substitution of formula B-10 in formula B-9 and than in the numerator of formula B-8 results in an unbiased estimator for EXP(μ) with minimum variance:

$$GM^{u} = GM * \phi\left(-\frac{s^2}{4}\right)$$

with

$$\Phi(t) = \sum_{k=0}^{\infty}\left(\frac{\Gamma_{\nu}}{\Gamma_{k,\nu}} * \frac{t^k}{k!}\right) \quad \text{B-11}$$

The accuracy of BASIC is such (see D.3) that formula B-11 is calculable for $GSD/M^2<6$. Because $M^2>=4$ (by definition) and GSD's are mostly smaller than 24 (see B.2.3), formula B-11 is effective in industrial hygiene practice. Depending on the values of s and M, O(t) needs up to thirty terms before a constant value is reached.

Contents - Index

B.2.3 Variance (1). The geometric standard deviation

The geometric standard deviation EXP(σ) is a parameter:

- of scale and shape (see Figure 21 in Shimizu 1988a p10, Shapiro 1991 p6),
- without a dimension,
- larger than unity (by definition).

If all sample data are equal, then the estimator of EXP(σ) equals unity and is useless.

According to Leidel (1977 page 73) group GSDs commonly occur in the range 1.5 to 2.5. In Dutch series Buringh (1989, 1991) found GSD values between 1.4 and 3.5. For the estimation of the sample size Corn (1985 p 176) uses a EXP(σ) between 2.2 and 2.5. Kromhout (1993) reports a median GSD=2.41 in 45 studies yielding almost 20000 measurements. GSDs seem to have been increasing the last few decades, especially in chemical industry (Kromhout 1993). A change in organizing work (from fixed post operations to team structure) is a possible explanation.

Some windows in this manual will show the value GSD=EXP(1)=2.71828 as an example.

B.2.3.1 The unbiased estimator

The unbiased estimator of EXP(σ) with minimum variance is the antilog of the standard deviation of the logarithms of the results:

$GSD = \mathrm{EXP}(s)$ B-13

with s as defined in formula B.4.

For confidence limits of the GSD see B.2.3.2

Contents - Index

B.2.3.2 Confidence limits

The variance ratio s^2/σ^2 follows, under the null hypothesis (s=σ), the Chi-square distribution (Abramowitz 1970 26.4). The limits, including EXP(σ) with confidence U%, are calculated using (Land 1988 page 98 4.1.2.):

$$GSD_{U\%} = GSD^{\pm \frac{\chi_{U\%,df}}{\sqrt{M}}} \qquad \text{B-14}$$

where $\chi_{U\%,df}$ is the square root of the Chi-square distribution with df degrees of freedom.

B.2.2.2 Confidence limits

LOG(GM)/μ follows, under the null hypothesis, the Student distribution (Snedecor 1980, 4.9). The upper and lower limits, including EXP(μ) with confidence U%, are calculated using (Land 1988 page 93 3.1 and 3.2):

$$GM_{U\%} = GM * GSD^{\frac{t_{U\%,df}}{\sqrt{N}}}$$ B-12

where $t_{U\%,df}$ is the deviate of the Student distribution with df degrees of freedom and U% is the desired percentage of the confidence limit, mostly 5% and/or 95%. The two-sided confidence interval is the fraction between two confidence limits.

Contents - Index

5 EXTRAPOLATION AND COMPLIANCE TESTING

The tabs <Upper Limit UTL> and <Mean UCL> calculate the effectiveness of workplace control measures on exposure by:

- extrapolating the exposure data to general results,
- testing exposure distribution against an industrial hygiene limit value.

Based on the exposure assessment goal and the required sample properties, Table 7 helps to choose the statistical method that is appropriate for extrapolation and compliance testing.

Table 7 Choosing the most appropriate method to generalize and test series sample exposure data.

Industrial Hygiene extrapolation and compliance testing deals with estimating:

- The true fraction of the distribution beneath a certain value
- The true value of a descriptive statistic, mostly the arithmetic mean

Remarks.
Calculations presented are performed exclusively with:

- the estimators of EXP(μ) and EXP(σ) and the number of degrees of freedom df,
- the industrial hygiene limit value H and the desired percentage U% ,
- the Noncentral Student or special cases of this distribution (Owen 1968).

The theory of the Lognormal methods used in this chapter is described extensively by Land (1988 page. 87-106).

Contents - Index

Effectivity, Effectiveness

The combined minimum of systematic (bias, accuracy) and random errors (efficiency, variance, precision). Increases if the systematic error of the first and second kind and the random error decrease

Contents - Index

Table 7 Choosing the most appropriate method to generalise and test series sample exposure data.

Exposure assessment goal	Required sample properties: Uncensored sample size	Lognormal-conformity	GSD²	Statistical method. Chapter (screen)
Long term $TWA_{reference\ period}$ control[3]: Which fraction of the $TWA_{ref\ per}$ exceeds the limit value $H_{ref\ per}$?	≥20	≥ very good	≤3.5	Standard Normal (Leidel) par. 5.1.2, (Screen 24)
	≥2	≥ acceptable	>1	Unbiased Student (Wilks) par. 5.1.3, (Screen 25)
What is the confidence δ that $H_{ref\ per}$ is exceeded ≤α% of the times?	≥2	≥ acceptable	>1	Noncentral Student (Tuggle) par. 5.1.4, (Screen 26)
Dose assessment. What is the: grab sample based chance of exceeding $H_{ref\ per}$?	df<25	≥ acceptable	≤3	Dose, UMP (Bar-Shalom) par. 5.2.1 (Screen 24)
$TWA_{ref\ per}$ based chance that dose $\beta_{ref\ per}$ exceeds $H_{ref\ per}$?	≥ 2	≥ acceptable	>1	Dose, approximate (Jahr) par. 5.2.2 Screen 27
Ceiling control: What is the confidence δ that $H_{ceiling}$ is exceeded ≤α% of the times?	≥5	Rejected	-	Nonparametric par. 5.3 -
	≥2	≥ acceptable	>1	Noncentral Student par. 5.3.3 Screen 26

Contents - Index

note 18

Statistical analysis can prove poor quality in a data series but cannot improve it (The "rubbish in, rubbish out" principle). The user should have basic knowledge on exposure assessment, including assessment strategy and compliance control and should act accordingly to perform extrapolation and compliance testing successfully.

Contents - Index

5.1.4 Noncentral Student, tolerance limits (Tuggle)

To improve the NIOSH decision scheme (Leidel 1977 figure 1.1 page 11) when GSD>2, the one-sided tolerance limit (OTL) method was introduced by Tuggle (1981, 1982). The Belgium Standard Institute (BIN 1987) officially published an assessment strategy based on Tuggle's method. Screen 27 displays the confidence δ that <α% of the exposures is over limit H.

The confidence $\delta(C\alpha_{<100-U\%}>H)$ with a fixed non-compliance probability α<100-U% is estimated with the following test statistic (based on Tuggle 1982 formula B-2 page 345):

$$T_{df,Z_\alpha} = \frac{\dfrac{\dfrac{LOG\left(\frac{H}{GM}\right)}{LOG(GSD)} - Z_\alpha * \sqrt{df}}{\dfrac{LOG(GSD)}{\sqrt{df}}}}{\dfrac{LOG(GSD)}{\sigma}}$$

formula 5-10

$T_{df,Z\alpha}$ is, under the null hypothesis, a deviate of the Noncentral Student distribution. The estimator $D\{C\alpha_{<100-U\%}>H\}$ is calculated using $T_{df,Z\alpha}$ the number of degrees of freedom df and the noncentrality parameter Zα, the deviate corresponding with the standard normal probability α=100-U%.

Input U, Tdf,Zα and/or df output D(Cα<100-U%>H)
αW50% - Not accepted U%=100-α value in screen 21
50-αW10-14% - central Student formula 5.6
UD99.9999% - Not accepted U%=100-α value in screen 21
0<Tdf,Zα<10-16 - 50 %
Tdf,ZαD10-16 W200 algorithms of Owen (1956 formula 3.9 & 1968 page 464)
Tdf,ZαD10-16 >200 Normal approximate Algorithm 26.7.10 (Abramowitz 1970)

According to Leidel (1977 page 118) an employer should try to attain δD95% confidence that workers exposure is with αW5% over the limit.

The upper or one-sided tolerance limit (UTL) $Ç\alpha_{<100-U\%,\delta D U\%}$ is estimated using:

$$C_{\alpha<100-U\%,\delta\geq U\%} = GM * GSD^{T_{df,z_\alpha,\delta}}$$

formula 5-11

The program calculates the one-sided upper tolerance limits (UTL) if the desired percentage U% is more than 50%. The UTL is the concentration under which U% of the population data lies with U% confidence.

The one-sided upper tolerance limit (UTL) is calculated as follows:

- first a linear interpolation from D=0% to D=100% with GM/GSD as the initiating value and GSD as the multiplication constant
- If the situation is reached that D(n+1)>100-U%>D(n), then an inverse linear approximation (General Falsi or False position; Abramowitz 1970 chapter 3.9.3) is performed. The accuracy of UTL is such that ABS[D(UTL)-U%)<0.01%.

If the desired confidence U% and the GSD are extreme high and the sample size is extreme low, then the interpolation may take some time.

A non-iterative algorithm to calculate the inverse Noncentral Student based $C_{\alpha<100-U\%,\delta}D_{U\%}$ does not exist. The approximations of Patel (1982 chapter 7.11 page 204 Cornish-Fisher expansion), Johnson and Kotz (1970 30.5 page 194) and Owen (1968 page 468 a comparison of different algorithms) are, for common industrial hygiene values of M (small <10) and GSD (large >2), inaccurate. One-sided tolerance factors $T_{df,\alpha,\delta}$ for specific values of $Z\alpha$ and $Z\delta$ can be found in the tables of Resnikoff (1957), Ciba-Geigy (1980), Odeh (1980), Bezemer (1981), Tuggle (1982), Hawkins (1991 Table IV.1 page 136) and the charts in BIN (1987). There are small differences between the tolerance factors used by Tuggle and Hawkins (both from NBS 1966) and the tables of Odeh (1980). The latest is considered more accurate.

Example 31 Noncentral student tolerance limits

5.1.3 Unbiased estimate of compliance (Wilks)

To improve the small sample bias of the NIOSH test statistic (see table 9) the unbiased method of Wilks was introduced in industrial hygiene (Scheffers 1987). The method has been developed in the quality control of large scale product lines with homogeneous conditions (e.g. chemical refineries, battery production; Proschan 1953 page 556), which is comparable with long-term non-compliance control in similar exposure groups. Screen 26 displays the unbiased estimation of the non-compliance probability and the corresponding tolerance limit.

The algorithm developed Wilks (1941) was proven mathematically correct by Proschan (1953). The unbiased non-compliance probability test against limit H is based on the following test statistic:

$$t_{df} = \frac{\mathrm{LOG}\left(\frac{\mathrm{H}}{\mathrm{GM}}\right)}{\mathrm{LOG(GSD)} * \sqrt{1 + \frac{1}{\mathrm{df+1}}}}$$ Formula 5-5

t_{df} follows, under the null hypothesis, the central student distribution. The right sided non-compliance probability A_t is the Student density distribution integrated from tdf to infinity:

$$A_{t_{df}} = \frac{\Gamma\left(\frac{df+1}{2}\right)}{\sqrt{df * \pi} * \Gamma\left(\frac{df}{2}\right)} * \int_{t_{df}}^{\infty} \left(1 + \frac{x^2}{2}\right)^{\frac{df+1}{2}} dx$$

with

$$t_{df} > 0$$

Formula 5-6

The Student distribution is a special case of the Noncentral Student (Owen 1968) with noncentrality parameter $Z_\delta=0$. non-compliance probability $A_{C>H}$ in 5.3 is calculated from t_{df} and df using:

ABS(tdf)	df	output AC>H %=100 * At
<1E-16	-	50 %
>=1E-16	<=30000	Power series of Owen (1968 page 465). Relative accuracy -A%/A%=1E-4
>=1E-16	>30000	Normal approximate Algorithm 26.7.8 (Abramowitz 1970)

The corresponding tolerance limit $Ç_{U>=100-\alpha\%}$ is derived from formula 5.5 by logarithmic transformation:

$$Ç_{U\%\geq 100-\alpha\%} = GM * GSD^{t_{\alpha,df} * \sqrt{1+\frac{1}{df+1}}}$$ Formula 5-7

The unbiased Student tolerance factor $t_{\alpha,df}$ is calculated from α and df using:

Input α % and df output tα,df

ABS(a-50)<0.00001 % ≥1 0

0.0001<= α <=99.9999% 1 $$\tan\left(\frac{(50-A)*\pi}{100}\right)$$ Formula 5-8

0.0001<= α <=99.9999% 2 $$X*\sqrt{\left(\frac{2}{1-X^2}\right)} \text{ with } X = 1 - \frac{A}{50}$$ Formula 5-9

0.0001<= α <=99.9999% >=3 and>=1000 Newton-Raphson iteration (Abramowitz 1970 3.9.5.)
on the power series of Owen (1968 page 465),
with the asymptotic expansion 26.7.5. (Abramowitz 1970)
as initiator. -t/t<1E-6

0.0001<= α <=99.9999% >1000 Asymptotic expansion 26.7.5. (Abramowitz 1970)

<0.0001% or >99.9999% Any value not accepted U% value in screen 21

The unbiased estimate is effective in large similar exposure groups. With small but efficient sampling plans the effectiveness of control measures in a SEG can be assessed using control charts, examining goodness-of-fit and non-compliance probability calculations, just as is done in quality control for over decades (see e.g. Morrison 1958). $Ç_{U\geq\#\#\%}$ overestimates $C_{U\geq\#\#\%}$ on the average. For an unbiased estimate $Ç_{UD\#\#\%}$, see Owen (1968 page 461).

Example 28 Wilks tolerance limits and compliance probability

Figure 15 Vinylchloride Control chart in a SEG PVC operators TWA8 hours over 1987-1993

Figure 16 Lognormal probability distribution of 37 VCM TWA8hours

Example 29 Unbiased estimate of the long-term TWA control effectiveness

Example 30 Different number of degrees of freedom

Contents - Index

Similar exposure group

- A group of workers performing the same tasks on the average and with a long term exposure variability that outshines any systematic between worker exposure variability (ad hoc).

Contents - Index

Example 28 Wilks tolerance limits and compliance probability

*.HYG name File description Size Units df H AC>H% ÇU<=5% ÇU>=95%

BOLEY62 TWA8 hour PAS 12 mg/m3 9 10 28.3 0.46 48.2
papermill total dust

COHEN132 maximum flood levels 20 106ft3/sec 18 - - 0.23 0.68

DEWELL24 TWA8 hour respirable 8 mg/m3 6 5 0.08 1.02 2.37
dust in foundry

DEWELL44 TWA8 hour PAS 11 mg/m3 9 5 89.8 3.82 32.4
welding [MIG] fume

GUPTA271 mice survival after 10 Days 7 - - 36.2 86.6
inoculation with
tuberculosis

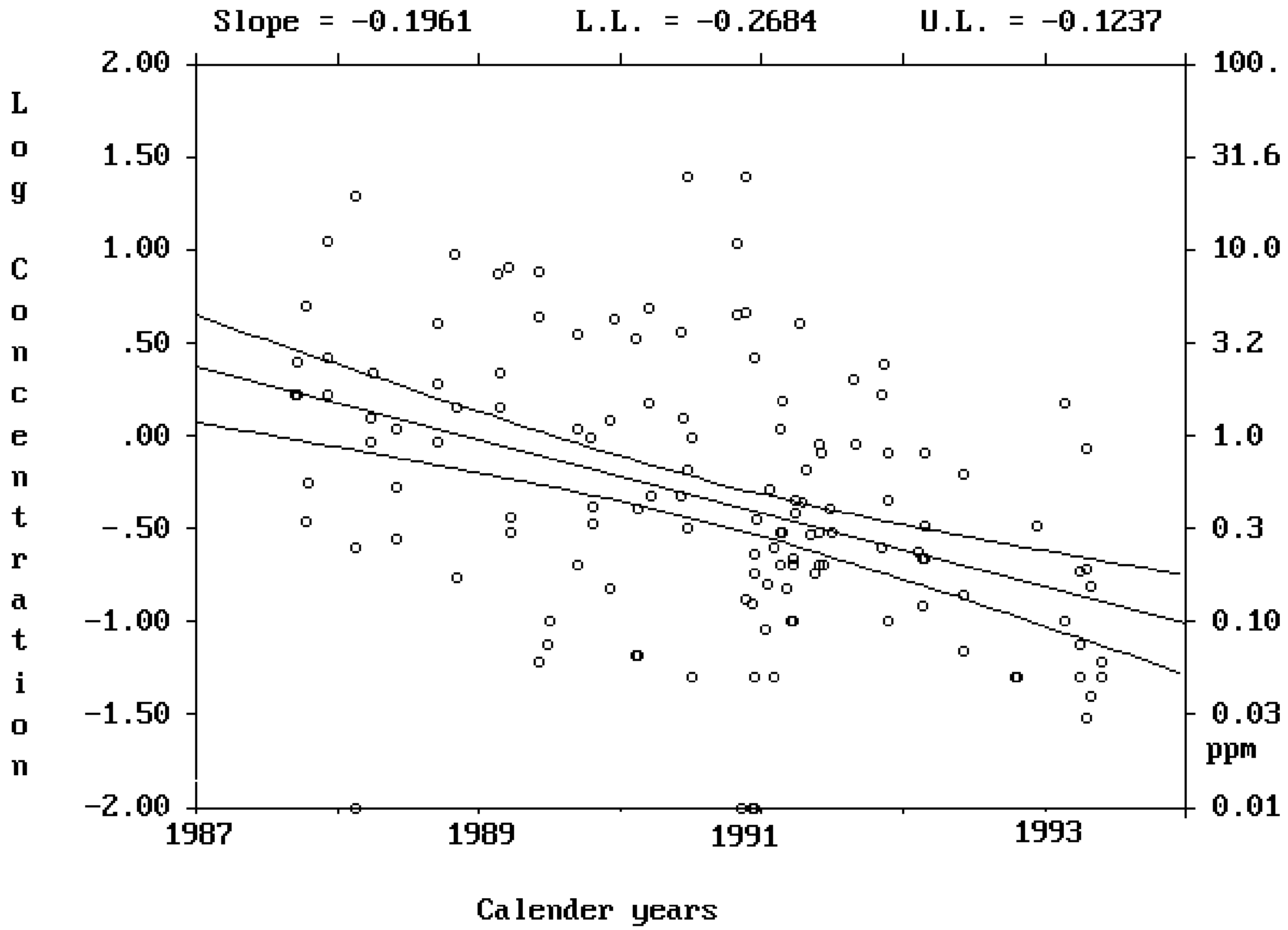
HAW117 TWA8 hour PAS 15 mg/m3 13 10 7*10-5 1.59 3.02
total dust

SCHNE224 Survival of items 50 TU 30 1 99.7 2.5 33.4
under stress

For SCHNE224 a lower quality limit is introduces of H=1 TU, which should be exceeded by 99% of the items.

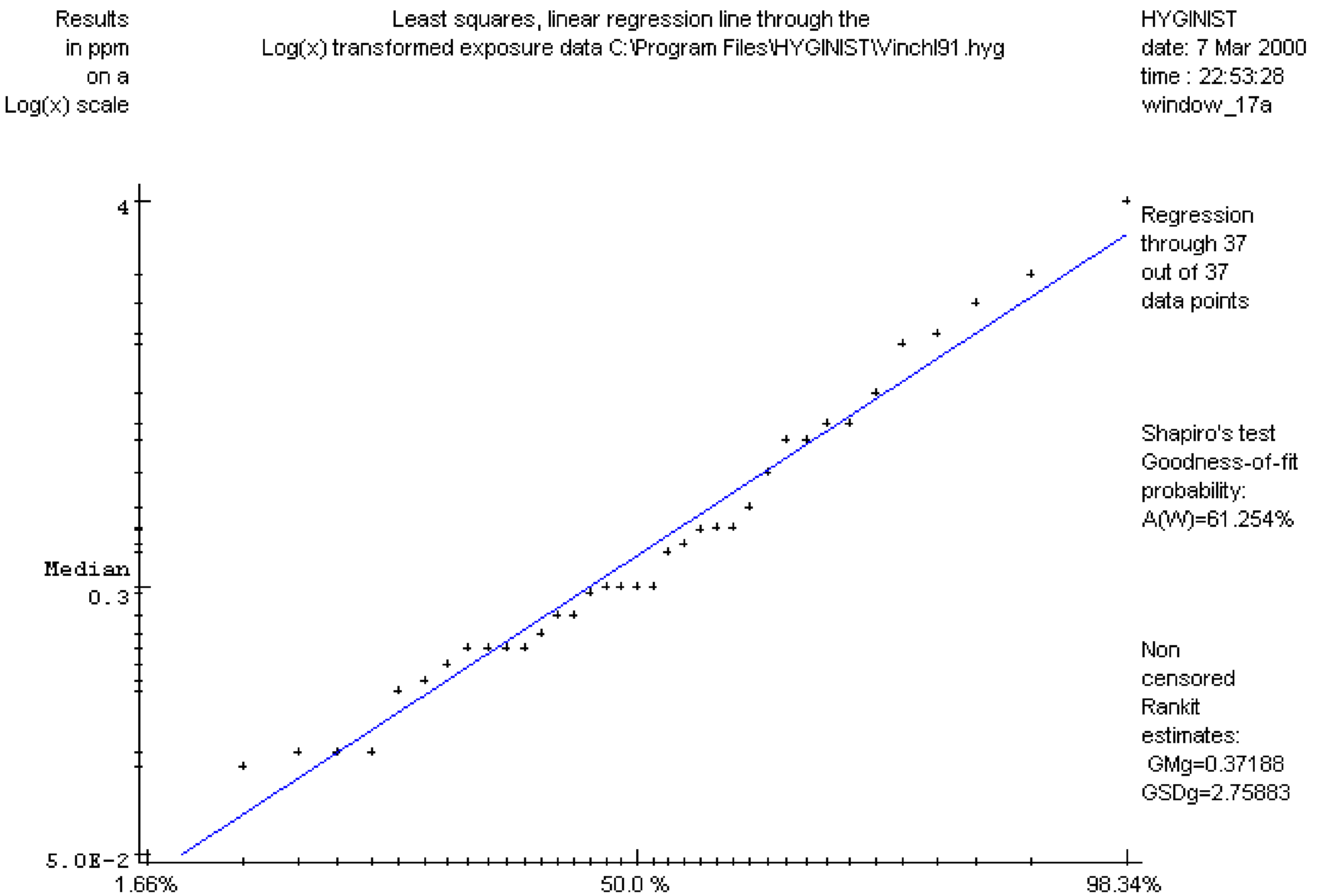
Contents - Index

Figure 15 Vinylchloride Control chart in a SEG PVC operators TWA8 hours over 1987-1993



Contents - Index

Figure 16 Lognormal probability distribution of 37 VCM TWA8hours



Contents - Index

Example 29 Unbiased estimate of the long-term TWA control effectiveness

*.HYG file Description

LEIDEL67 The 10 TWA8 hour Dioxane, see Figure 1 and screen 25), sampled on different days in a period of 6 months on one employee are tested against the H8 hour=100 PPM. Because GSD=1.63 and LOG(H/GM)=0.24 are both small, the bias of the standard normal method is also small (see table 9). The unbiased estimate AC>H=32.4% corresponds with the Pn=.0309 calculated by Leidel (1977 page 69).

LEIDL104 With 24 TWA's in the job category "Mix man" in a facility using Methyl methacrylate the unbiased estimate of the non-compliance probability is $A_{C>H}$= 5.7068 %, indicating (in deviation with the NIOSH method see Example 27) that control measures are necessary.

VINCHL91 In a PVC production plant, the SEG polymerisation shift workers are monitored on Vinylchloride using a routine plan using a stratified PAS TWA8 hour sampling plan (every four weeks, one day, three consecutive measurements, around the clock). The results since 1987 are displayed in the exposure control chart Figure 15. A limit $H_{8\ hour}$=7 PPM is derived from the European Community exposure limit $H_{EC,1\ year}$=3 PPM (EC 1978 L197/16). The Lognormal fit of the 37 TWA8 hour data in 1991 is complete (Figure 16 and $A_{W=0.974}$=63 %). With df=36, GM=0.37, GSD=2.73 unbiased estimate is $A_{C>H}$=0.3 % and the upper tolerance limit $_{ÇU>=95\%}$=2.1 PPM, indicating that the control measures are effective. See Example 34 for the test against $H_{1\ year}$=3 PPM.

- In a sample with GSD=2.71828 the ratio upper tolerance limit/geometric mean is for: df=1 -> C95%/EXP(μ)>=2182.

So compliance is reached if the geometric mean exposure EXP(μ) is below 0.05% of the limit value H. The lower accuracy limit of measurement methods for contaminants in workplace air will in most cases not include H/2000. Increasing df to infinite decreases C95%/EXP(μ) asymptotically to the standard Normal ratio of 5.18 (see further chapter 7).

Contents - Index

Example 30 Different number of degrees of freedom

*.HYG file Description

SCHNE224 The survival of items under stress is Lognormal distributed (Schneider 1986 page
224; see Figure 5). If the (accelerated) lifetime is above 1 time units (TU), then an item conforms the quality limit. 15 items out of a sample of 50 failed before 6 TU.
In a conservative approach, using df=M'-1=14, AC>1 TU=99.4% and Ç5%=2.3 TU, indicates only a borderline difference from the values calculated in Example 29.

Contents - Index

note 21

The tables 1.#.# of Odeh (1980 page 18 through 69 using γ, 1-P and N for confidence γ, non-compliance probability α, and sample size M respectively) provide tolerance factors for:
δ= .995, .99, .975, .95, .9, .75, .5, .25, .1, .05, .01, .005
1-α= .75, .90, .95, .975, .99, .999
M= 2(1)100(2)180(5)300(10)400(25)650(50)1000, 1500, 2000, 3000,5000,10000, ∞

Contents - Index

Example 31 Noncentral student tolerance limits

*.HYG file Description
LEIDL104 The 24 TWA$_{8\ hours}$ indicate that the confidence of being in compliance for job category "Mix man" in a facility using Methyl methacrylate is δ=51%.
Thus consecutive sampling plans with M=24,
will have a chance of 51% of finding non-compliance probability αW5%.

SOLV198 From 45 TWA$_{8\ hours}$ Hydrocarbons during solvent based paint rolling and spraying the two-sided 95% tolerance limits with δ=95% confidence (9.4 and 1076 mg/m3) were calculated using the tolerance factor k7=2.408 from Ciba-Geigy (1980).
The one-sided tolerance factor for calculating the 2.5% and 97.5% limits with δ=95% is $T_{df,\alpha,\delta}$=2.463 and provides 8.9-1141 mg/m3.

- Tuggle (1981 p 491) evaluated 3 measurements of 0.60, 0.58 and .63 with a PEL of 1.0.
The sample 10log=-0.22 (GM=0.6) and
s=0.02 (in real s=0.018 thus GSD=1.02).
With the use of the tolerance factor Tdf,α=5,δ=95=7.655 the
UTL=-0.22+7.655*.22=-0.08 is less than $_{10}$log(PEL)=0.
Only because of the unrealistic small sample GSD
this small sample size problem is in compliance.

SCHNE70 The 97.5% UCL of U=90% of the 96 locomotives,
With 37 failures before 135,000 Miles is estimated from the data presented by Schneider (1986 example 6.3 page 202) as 547,016 Miles.
This mileage value corresponds, however, with
$D_{C\{\alpha<10\%\}}$>547,016=90.5% for the Noncentral Student with df=65.
The tolerance limit calculated with table 1.4.1. of Odeh (1980) is
$Ç\alpha_{<10\%,}\delta D_{97.5\%,df=65}$=174,249*2.11718$^{1.657}$=603,883 Miles.

HAW117 From 14 random TWA8 hour PAS total dust above LL=1.4 mg/m3 the one-sided upper tolerance limit including 90% of the concentrations with 95% confidence and df=13 is H=3.1712 mg/m3.
This is comparable with the values calculated by Hawkins (1991) page 140 table IV.2 bottom line (UTL_c=3.06 and 3.29). See also Example 57.

- In a sample with GSDD2.71828 the ratio upper tolerance limit/geometric mean is for:
- df=1 º $C\delta D_{95\%,\alpha<5\%,}$df=1/EXP($\mu$)D2.5384*10^{11}
- df=2 º $C\delta D_{95\%,\alpha<5\%,}$df=2/EXP($\mu$)D2113.28

Due to the limitations in the industrial hygiene sampling/analytical methods, a concentration of H/2000 is difficult to assess for most contaminants in workplace air. Increasing df decreases $C\delta D_{95\%,\alpha<5\%}/EXP(\mu)$ asymptotically to 5.18.

Contents - Index

5.1.2 Standard Normal (Leidel)

This generally accepted, long-term non-compliance probability method obtains its status from the recommendation of the USA's "National Institute of Occupational Safety and Health" (Leidel 1977). As is displayed in screen 25, its application is tied to such rigorous criteria on sample size, Lognormal fit and variance, that most industrial hygiene data sets cannot be evaluated using this method.

To estimate $\alpha_{C>H}$, Leidel (1977 page 69) replaced log(C)-μ and σ in formula 5-1 with the estimators LOG(H/GM) and LOG(GSD) respectively. The non-compliance probability test is based on the following statistic:

$$Z = \frac{\mathrm{LOG}\left(\frac{H}{GM}\right)}{\mathrm{LOG}(GSD)}$$ formula 5-2

Z is, under the null hypothesis, a deviate of the standard Normal distribution with probability $A_{C>H}$:

$$A_Z = \frac{1}{\sqrt{2 * \pi}} * \int_Z^{\infty} e^{-x^2/2}\, dx$$ formula 5-3

The standard Normal distribution is a special case of the Noncentral Student (Owen 1968) with df=infinite and noncentrality parameter Zδ=0. non-compliance probability $A_{C>H}$ in 5.3 is calculated from Z using:

input value Z output $A_{C>H}$=100 * A_Z

ABS(Z)<1E-14 50 %
ABS(Z)<+5 the power series 26.2.11 of Abramowitz (1970). Accuracy -A<1E-12%
5<ABS(Z)<=37 the approximate fraction expansion AS 66 of Hill (1973). Accuracy -A<1E-11%
<-8 100 %
-8<=Z<37 1E-300 <=A_Z<+99.9999999999999 %
>37 0 %

The tolerance limit ÇUD$_{100-\alpha\%}$ is derived from formula 5.2, by exponential transformation:

$$\text{Ç}_{U\geq 100\alpha\%} = GM * GSD^{Z_\alpha}$$ formula 5-4

The standard Normal tolerance factor Zα is calculated from α using:

Value of α output Zα

ABS(α-50)<1E-4 % 0

D1E-4% and <=99.9999% Polynomials AS 111 of Beasley (1977), followed by the inverse interpolation of Abramowitz1970 p954 Example 5).

Accuracy -Z<5*1E-16

<1E-4 % or >99.9999% Not accepted U% values in screen 21

Example 26 NIOSH tolerance limits and compliance probability

Example 27 Calculating standard normal tolerance limits from literature sources

Contents - Index

Example 26 NIOSH tolerance limits and compliance probability

*.HYG name File description Size units df H AC>H % ÇUW5% ÇUD95%

LEIDEL67 TWA8 hour Dioxane 10 PPM 9 100 31.0 35.1 175.4
LEIDL104 TWA8 hour Methyl methacrylate 24 PPM 23 100 4.7 12.1 98.0
RANDOM20 standard Lognormal deviates 2000 - 1999 - - 0.19 5.11
SCHNE70 Locomotive controls 96 103 miles 37 - - 50.7 598.4

See Example 28 for more details.

Contents - Index

Example 27 Calculating standard normal tolerance limits from literature sources

*.HYG file description

LEIDL104 With 24 $TWA_{8\ hour}$ Methyl methacrylate in the job category "Mix man", the NIOSH non-compliance probability is $A_{C>H}$=4.7 % (see screen 25 and Example 26).
Engineering controls are not necessary (Leidel 1977 p104).
This conclusion contrasts with the unbiased method (see Example 29).

LEIDEL67 Leidel (1977 4.4 p65) utilised 10 $TWA_{8\ hour}$ Dioxane (see Figure 1) to illustrate the NIOSH method of using the probability of non-compliance to decide if engineering controls should be installed. The non-compliance probability $A_{C>H}$=31% equals P_n=.0309 calculated by Leidel (1977 page 69).
Since non-compliance is so obvious, there is no reason to criticise the use of the standard normal method because of its small sample size (see table 7).

RANDOM20 The 5th and 95th percentile values in the series 2000 Monte Carlo standard Lognormal deviates are 0.1933 and 5.053, and agree with the calculated values .1927 of 5.110 (see Example 26)

SCHNE70 The Lognormal fit (see Figure 6) and the number of failed locomotives are such that standard Normal extrapolation is allowed (see table 4). Schneider (1986 example 6.3) calculates the 10th percentile as $10^{1.826+3}$=66,988 Miles using the bias corrected estimators of μ and σ. The standard Normal lower tolerance limit is almost equal: $ÇUW_{10\%}$=66,634 Miles.

Contents - Index

Table 4 Lognormal goodness-of-fit decision scheme

What's the shape of logarithmic Is the Goodness- Transfor- Uncen Lognormal
transformed scatterplot sample of-fit mation sored conformity
within the accuracy range, censored A(W)% fit Sample
along the regression line ? ? size

straight line no ≥95 best ≥20 perfect
ditto yes - best ≥20 very good

slightly curvilinear no ≤95 best ≥20 good

linear tendency, slightly oscillating no ≥95 best ≥3 good
ditto no 5-95 ≥2nd best ≥3 ≥acceptable
ditto yes - ≥2nd best ≥3 ≥acceptable

unknown (descriptive statistics only),
but reliable assessment strategy - - - ≥2 Acceptable

Curvilinear no <95 <2nd best ≥3

Oscillating no <5 - ≥3 Lognormality

oscillating or curvilinear yes - <2nd best ≥3 Rejected

Contents - Index

note 17

This categorical classification increases from rejected (nonconformity), acceptable, good, very good, to perfect (complete conformity).

Contents - Index

note 14

Low A(W) values of the omnibus W-test of Shapiro (1965) indicate low conformity with the Lognormal shape.

note 15

The Lognormal goodness-of-fit rank order, among the transformations (**<** less than, **<=** at most, **=>** at least).

Contents - Index

note 16

The number of exposure data M` within the accuracy range

Contents - Index

note 20

See § 5.1.1 for the differences between the three statistical methods

Contents - Index

---

note 19

If the measurement method random error (CV_t*AM) or the grouping interval (@C) are larger than the arithmetic standard deviation w (see § 4.1.1), then GSD should be adjusted (see §B.6). If GSD>1.4, then random error is in most cases irrelevant.

B.6 GSD and CVt

Part of the sample variance may be explained by the non-systematic measurement error CV_t. Based on the relation between EXP(σ) and the arithmetic mean β and variance ω^2 (Leidel 1977, table M-1):

$$\sigma = \sqrt{\mathrm{LOG}\left(1 + \frac{\omega^2}{\beta^2}\right)}$$

with the population descriptiv e statistics :

σ = standard deviation of log (C_i)

ω = arithmetic standard deviation of C_i

β = arithmetic mean of C_i

A relative adjustment factor R can be derived using $w^2 = w_t^2 - s_e^2$ with:

- w_t^2 the total arithmetic variance of the sample data (see formula B-16),
- s_e^2 the variance caused by measurement error and, according to Leidel (1977 p50 NOTE), $s_e^2 = (CV_t * H)^2$.

$$GSD_{adj} = R_{adj}^{GSD} * GSD$$

with

$$R_{adj}^{GSD} = \frac{\mathrm{EXP}\sqrt{\mathrm{LOG}\left(1 + \frac{w_t^2 - (CV_t * H)^2}{AM^2}\right)}}{\mathrm{EXP}\sqrt{\mathrm{LOG}\left(1 + \frac{w_t^2}{AM^2}\right)}}$$

AM and w_t^2 are the unbiased estimators of the arithmetic mean β and the arithmetic standard deviation ω (see formulas 5.16 and 5.18). If $w_t < CV_t * H$, then the square root of a negative number is extracted in the denominator and formula 5.27 becomes incalculable. In Example 56 values of GSD_{adj} and w_{adj}/AM_{adj} are calculated for relevant values of GSD, M, and CV_t. For M>10 or GSD>2.71828 the influence of random measurement errors CV_t<.35 is irrelevant. The same applies for MD2, GSDD1.4 and CV_t<.1.

Example 56 Adjusting GSDs and w/AM for different values of M and CV_t

[Example 57 CVt influence on GSD in example data]

Contents - Index

Example 56 Adjusting GSDs and w/AM for different values of M and CVt

Sample size	Sample descriptive					
Statistics			CV_t=.05		CV_t=.25	
M	GSD	w/AM	GSD_{adj}	w_{adj}/AM_{adj}	GSD_{adj}	w_{adj}/AM_{adj}
2	1.1	0.0952	1.0845	0.081	-	-
2	1.4	0.3303	1.3951	0.327	1.2562	0.223
2	2	0.6425	1.9970	0.641	1.9219	0.603
2	2.71828	0.8611	2.7157	0.860	2.6513	0.836
2	3.5	1.0024	3.4974	1.002	3.4334	0.984
5	1.1	0.0953	1.0846	0.081	-	-
5	1.4	0.3314	1.3952	0.331	1.2600	0.224
5	2	0.6822	1.9972	0.680	1.9287	0.636
5	2.71828	0.9679	2.7161	0.967	2.6633	0.936
5	3.5	1.1915	3.4980	1.191	3.4513	1.166
10	1.1	0.0954	1.0846	0.081	-	-
10	1.4	0.3395	1.3952	0.336	1.2623	0.228
10	2	0.6390	1.9974	0.717	1.9342	0.669
10	2.71828	1.0736	2.7164	1.072	2.6724	1.038
10	3.5	1.3907	3.4985	1.390	3.4637	1.363
50	1.1	0.0954	1.0846	0.081	-	-
50	1.4	0.3446	1.3953	0.341	1.2654	0.231
50	2	0.7680	1.9976	0.766	1.9407	0.714
50	2.71828	1.2387	2.7169	1.239	2.6831	1.202
50	3.5	1.7604	3.4991	1.759	3.4774	1.732

The adjustment of GSD and w/AM>CV_t, increases if w/AM is about equal to CV_t.

Contents - Index

5.1 Long-term TWA control

To assess long-term compliance with hygiene limit H, 3 statistical extrapolation methods are presented and ranked in the increasing probability that H is considered as being an element of the exposure distribution. Use table 7 to choose the appropriate method.

Table 8 combines the names found in literature for measures of long-term TWA extrapolation parameters, and the names (in bold) and symbols used in the screens 25 through 27.

Table 8 Names, parameters and estimators for assessing long-term TWA control

Upper and lower limits Ç $\delta_{(100-U)/2<=\#\#\%}$ and Ç$\delta_{(100+U)/2>=\#\#\%}$ of the two-sided tolerance interval U% are calculated by introducing successively $\delta_{(100-U\%)/2}$ and $\delta_{(100+U\%)/2}$.

Contents - Index

Table 8 Names, parameters and estimators for assessing long-term TWA control

Name Population sample chapter
parameter estimate
Long-term non-compliance probability (NIOSH 1977 p65) αC>H AC>H 5.1.2 &
Percentage of the TWAs above the limit H 5.1.3
One-sided or right-sided fraction
Fraction TWAs in compliance with the limit value H
Chance αC>H that one TWAreference period exceeds H
Acute health hazard (Hawkins 1991 page 56)

Upper tolerance limit
One-sided or right-sided tolerance limit CUD##% ÇUD##%
Concentration above U% of the population 5.1.2 &
Lower tolerance limit 5.1.3
One-sided or left-sided tolerance limit CUW##% ÇUW##%
Concentration below 100-α% of the population

Confidence δ of desired α δ(Cα<100-U>H) D(Cα<100-U>H) 5.1.4
Confidence δ that <α% of the concentration are over H
non-compliance probability α with confidence δ α(CδDU%>H) A(CUD##%>H) 5.1.4
Fraction TWAs with confidence δ in compliance with H
Chance α with confidence δ that one TWAref per exceeds H

Upper tolerance limit with confidence δ Cα<100-U,UD##% Çα<100-U,UD##% 5.1.4
One-sided tolerance limit (OTL, Tuggle 1982)
Concentration above 100-α% of the population with confidence δ

Contents - Index

5.2 Dose assessment

The Tab "UCL" helps to assess the:
- $TWA_{ref\ per}$ or the chance of exceeding $H_{ref\ per}$, from a series grab samples,
- the Lognormal arithmetic mean $\beta_{long\ term}$ and the chance that it exceeds $H_{ref\ per}$, from a series $TWA_{ref\ per}$,
- the chance that the mean of a series $TWA_{ref\ per}$ exceeds $H_{ref\ per}$, on the average.

This is useful for:

- the "Classification of exposure for an 8-hour TWA standard, based on grab or partial period, consecutive sample measurements" (Leidel 1977 page 55) and
- the cumulative or average daily dose (Seixas 1988) or the chronic health hazard (Hawkins 1991).

Table 10 displays the symbols and some of the names found in literature for the extrapolation parameters of the Lognormal arithmetic mean and the names (in bold) used in screen 24.

Table 10 Names, parameters and estimators of dose assessment

Upper and lower confidence limits Çβ, (100-U)/2=<% and Çβ, (100+U)/2##% of the arithmetic mean are calculated by introducing successively (100-U%)/2 and (100+U%)/2 in screen 21 .

Contents - Index

Table 10 Names, parameters and estimators of dose assessment

Name Population Sample
Parameter Estimate

Probability α that mean β exceeds limit H $\alpha_{\beta>H}$ $A_{\beta>H}$
non-compliance probability α of the arithmetic mean β.

Upper Confidence limit of β $C_{\beta,U>=\#\#\%}$ $Ç_{\beta,U>=\#\#\%}$
One-sided or right sided confidence limit of the mean
Upper concentration with confidence 100-α% over population mean β

Lower Confidence limit of β $C_{\beta,U<=\#\#\%}$ $Ç_{\beta,U<=\#\#\%}$
One-sided or left sided confidence limit of the mean
Lower concentration with confidence α% over population mean β

7.1 Stop sampling

Tab 34a displays the situation when $A_{t,df=1}<\alpha$, so in which compliance is reached for every sample size M>1.

In this special case the Wilks test statistic is:

$$Az \le 100\text{-}U\% \le A_{t_{df=1}}$$ formula 7-4

Because the Student distribution is only defined for integer values of df, the confidence of compliance with the limit value is higher than expected on the average even for df=1. In statistical terms it is not possible to conclude "Stop measurement program, the confidence is sufficient!". This is the domain of the experienced industrial hygienist, but this situation could justify such a decision.

From screen 34a it can be concluded that if:
- the ratio GM/H<.25,
- the variance GSDW1.20 and
- the desired noncompliance probability α=5%

then the exposure will be considered as in compliance on the average for every sample size. The upper line in table 1 of Scheffers (1987) provides, for different values of GSD, the ratio GM/H for which sampling is less useful.

Contents - Index

Example 50 The most effective measures to quantify exposure

Which descriptive statistics are used depends on the agent, the goal of the strategy, the desired extrapolation, the shape of the distribution and the limit value. Extremes are effective when dealing with Ceiling limits or when agents have an acute toxic response over a specific concentration (cyanide, ammonia and similar agents). The arithmetic mean is an effective measure for the mean mass for cumulating agents (asbestos, quartz and similar agents). Geometric mean and variance are essential for all extrapolation and inference statistics. Log(GM) is being used in noise exposure (in dB(A))

Contents - Index

Example 53 Performance of estimators of censored sample descriptive statistics

Three out of ten mice inoculated with tuberculosis did not die within 60 days (GUPTA271). Schneider (1988 p69) used this Lognormal distributed data set to illustrate the performance of different estimators of μ and σ in small censored samples. The results of the rankit estimator was added.

Name of the method page in symbol Estimators
Schneider (1986) GM in days GSD
Ignoring censored data - - 50.7 1.16
Censored data at 65 days - - 52.8 1.17
Best linear unbiased 85 BLU 55.2 1.23
Maximum likelihood 69 ML 55.2 1.20
Restricted ML 103 RML 55.4 1.21
Dixon's 88 - 55.5 1.23
Modified maximum likelihood 104 MML 55.2 1.20
Bias corrected ML 110 ML_c 55.6 1.23
Rankit regression 80 LU 56.0 1.24

Using only the non-censored data ($=1/7*\sigma x_j$ and $s=\{\sigma(x_j-)^2/6\}^{+1/2}$) or a fixed value for the censored data (Hornung 1990, Hawkins 1991 p104) underestimate both EXP(μ) and EXP(σ) rather extreme. The Rankit regression estimators result in the highest values for GM_g and GSD_g. This is suitable for industrial hygiene purposes because of the low chance of underestimating the location and variance of the exposure

Confidence d of one sided tolerance

(1-δ) the probability of accepting the desired compliance probability correctly (Tuggle 1982)

Confidence (in relation to a fraction)

The chance that the tolerance interval contains the desired fraction of the population

Contents - Index

Confidence (in relation to the parameter)

Confidence (in relation to the parameter), the chance that the confidence interval contains the parameter of estimation

---

Consistent

- Consistent estimate, see Asymptotic unbiased estimator

Contents - Index

H Terms and definitions

The explanations are based on Kendall (1982) or subsequently the most suitable term in literature. References between ().

- Accuracy, the difference between the measurement and the truth (Hawkins 1991), the extent in which calculations or estimations, the true value approximate
- Asymptotic unbiased or consistent estimator, an estimator which expectation equals the parameter by infinite sample size
- Autocorrelation, the phenomenon that observations in a series are dependent upon the preceding
- Averaging time, a period of time for which the measuring procedure yields a single result (prEN 482)
- Bias, systematic error (Hawkins 1991 156). Consistent deviation of the results from the true value (ISO 6879)
- Censored distribution, a distribution with a known fraction of observations outside the detection range (Hald 1949)
- Confidence δ of one sided tolerance; (1-δ) the probability of accepting the desired compliance probability correctly (Tuggle 1982)
- Confidence (in relation to a fraction), the chance that the tolerance interval contains the desired fraction of the population
- Confidence (in relation to the parameter), the chance that the confidence interval contains the parameter of estimation
- Consistent estimate, see Asymptotic unbiased estimator
- Degrees of freedom, the mutually independent, number of elements in a sample from the distribution
- Descriptive statistic, a number representing a probability distribution
- Dispersion, variance (Sarhan)
- Effectivity, Effectiveness the combined minimum of systematic (bias, accuracy) and random errors (efficiency, variance, precision). Increases if the systematic error of the first and second kind and the random error decrease
- Efficiency, according to Fisher (1921) an estimator is efficient if his population variance is the smallest (Kendall)
- Error of the second kind, accepting an invalid hypothesis. Also specificity (rejecting compliance)
- Error of the first kind, rejecting a valid hypothesis. Also sensitivity (detecting non-compliance)
- Error, see effectivity
- Estimate, the value of the estimator (Kendall)
- Estimator, a rule or method to estimate a constant in the population (Kendall)
- Expectation, the arithmetic mean of a probability distribution
- Extrapolation, to estimate the value a variable outside its tabulated or observed range
- Fraction, the number of elements of a population with a given property, divided by the size of the population
- Similar exposure group, a group of persons performing the same tasks and with a random exposure variability which overshadows the systematic in-between person variability (ad hoc).
- A group of workers with identical probabilities of exposure to a single environmental agent (Hawkins 1991 page 5).
- A group of employees who experience agent exposures similar enough that monitoring of any worker in the group provides data useful for predicting exposures of the remaining workers (Hawkins 1991 page 160).
- Inference, the process of deriving from assumed premises either the strict logical conclusion or one that is to some degree probable (The Random House college dictionary)

- Interpolation, to insert, estimate or find an intermediate term (in a sequence)
- Kurtosis, the fourth moment
- Location, mean (Sarhan)
- Lognormal descriptive statistics, location and variance parameters describing a Lognormal distribution
- Maximum likelihood, one of the methods that provides estimators of the parameters of a parent population from samples. ML estimators fulfil certain criteria for consistency, efficiency and sufficiency. In fact ML estimators are consistent, tend to Normality for large sample size, have minimum variance in the limit at least, and provide sufficient statistics where such exist (Kendall 1947 volume II par.. 17.22)
- Median, the most central result or, in case that sample size is even, the arithmetic mean of the untransformed two most central data.
- Minimum variance estimate, a method of calculating an estimator with the characteristic of maximum efficiency
- Modus, the value in a distribution with the highest frequency.
- Monomorphic group, A group of workers whose individual mean exposures compromise a single log-normal distribution. Term introduced by Rappaport (1991 page 66) but should be avoided. See Similar exposure group.
- Mutually independent, in opposite with autocorrelation
- non-compliance probability, the population fraction measurements over the industrial hygiene limit value
- Occupational exposure assessment, comparing exposure with the limit value (prEN 689)
- Omnibus, term used by d'Agostino (1971) to characterise tests on shape that combine aspects of both skewness and kurtosis
- Performance, general requirements on the effectiveness of the exposure assessment (prEN 482)
- Periodic measurements, the regular check if exposure conditions have changed (prEN 689)
- Population, collection of elements on which the conclusions of the statistical evaluation are related to
- Power, the chance of rejecting a unvalued null hypothesis
- Precision, random error
- Precision, the size of the deviation from the mean of the observations (Hawkins 1991 163) the closeness of agreement between the results obtained by applying the method several times under prescribed conditions (prEN 482)
- Random drawing, taking an element from a population with a method that is independent of all properties of the element
- Range of accuracy, the values for which the measurements are considered to be reliable
- Reliability, see confidence
- Robust, the property of a test that it works well for a wide variety of population types (Snedecor 1980 p 135)
- Rankit, the expectation of the standard normal ordering (Fisher 1938)
- Reference period, the specified period of time stated for the limit value of a specific agent (prEN 689)
- Sample, a series drawings from a population
- Scale, variation
- Selectivity, degree of independence from interferents (prEN 482)
- Skewed distribution, an asymmetrical distribution (Kendall)
- Sustainment, the capacity of individuals or groups to bear exposure
- Truncated distribution, a distribution with an unknown proportion result outside the measurement reach (Hald 1949)
- Upper tolerance limit, the upper boundary containing at least the desired fraction of the population
- Unbiased estimator, an estimator with the characteristic that the expectation equals the parameter for every sample size
- Workplace, the defined area or areas in which the work activities are carried out.

Degrees of freedom

The mutually independent, number of elements in a sample from the distribution.

Contents - Index

Degrees of freedom

The number of degrees of freedom for the censored and non-censored case are calculated using:

```
If DetecNumAantal < M Then
    Nu = (M + DetecNumAantal)  2 - 1
    If Nu > 2 * DetecNumAantal Then Nu = 2 * DetecNumAantal
Else
    Nu = M - 1
End If
```

With:
M is sample size
DetecNumAantal is the uncensored number of measurements
Nu is the Number of degrees of freedom

Contents - Index

Descriptive statistic

A number representing a probability distribution

Contents - Index

Dispersion

Variance (Sarhan)

B.1 Goodness-of-fit

Following Geary (1947) one could say that "Lognormality is a myth. There never was and will never be, a Lognormal distribution.". Exposure data are only best represented by the Lognormal model (see B.4.1). Shapiro (1990 page 5) states that the reason to study the shape of a distribution is ".... whether or not it is reasonable to approximate the data with the model, not whether the data came from the hypothesized distribution". Thus studying goodness-of-fit of industrial hygiene exposure data is functional (Hawkins 1991 page 58), it can be done in samples of at least 3 exposure data and in two different ways:

- graphically, with a probability plot (B.1.1),
- quantitatively, with a test on Lognormality (B.1.2).

For both it is true that the effectiveness increases with sample size. The interested reader in "distributional assumption testing" is referred to Shapiro (1990).

B.4.1 General results

df equals the number of sampling periods minus unity (M-1) if:

- the total exposure duration ? is much larger (up to infinity) than the total sampled time (?>>t*M, t= sampling duration), the TWA situation,
- the number of consecutive exposure periods $M\tau$ is much larger (up to infinity) than the number of the sampled periods M ($M\tau >> M$), in the long-term compliance control situation.

Most exposure assessment strategies use df=M-1 and thus extrapolate implicit to general results.

B.1.1 Graphical

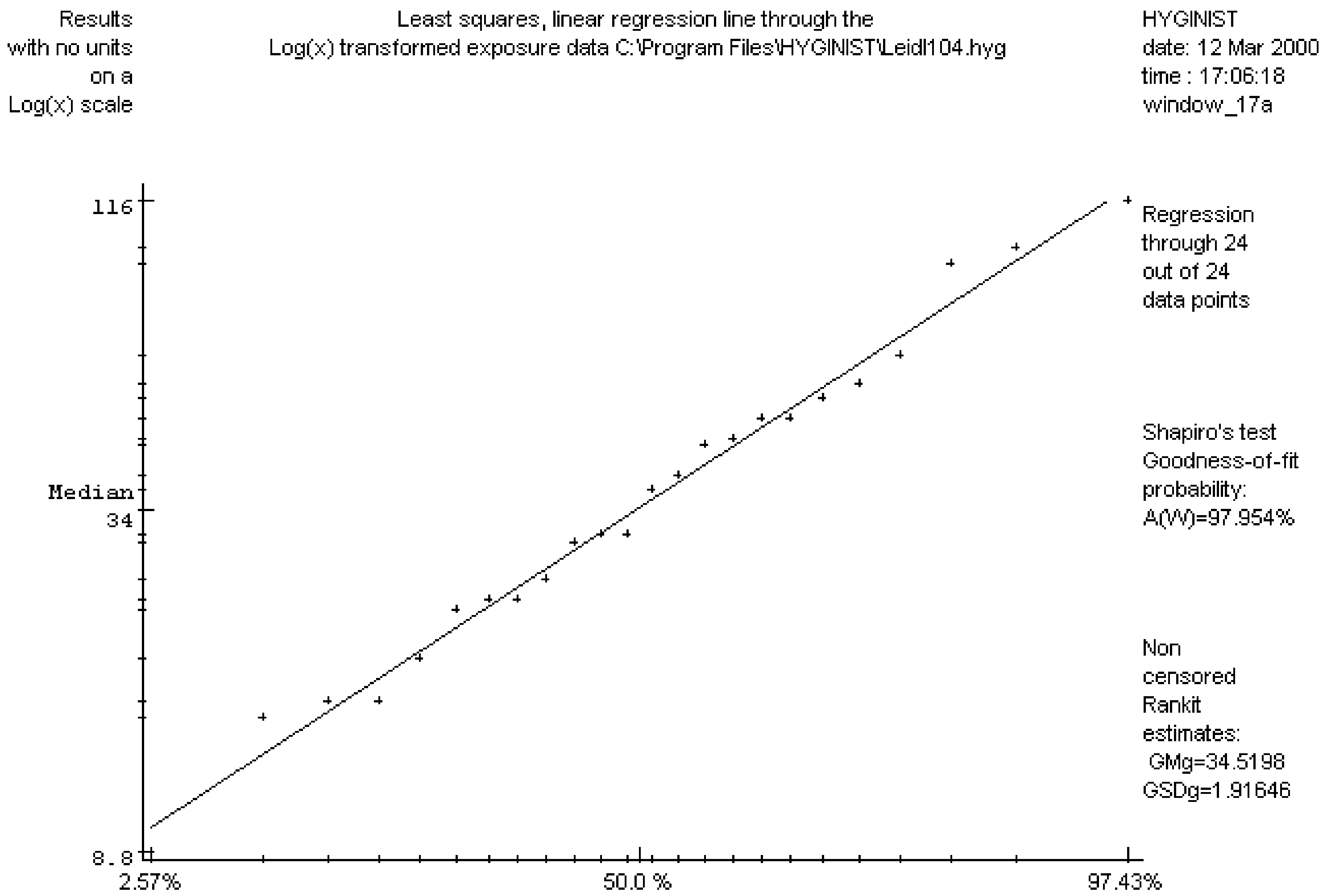
The examination of Lognormal shape with probability paper has a long tradition in industrial hygiene (Oldham 1953, Leidel 1977 page 102) and other quality control situations (Morrison 1958, Snedecor 1980 4.13). In probability paper the standard normal deviates or their probabilities (the expected values) are located on the horizontal axis and the logarithmically transformed, observed exposure data c_i in ascending order are located on the vertical axis. In the ideal case the crosshairs form a straight line through the (non existing) origin (see Figure 22. The decision that the sample data are from a (Log)normal distribution is subjective in this case.

Figure 22 Lognormal probability plot on 24 $TWA_{8\ hours}$ methyl methacrylate (LEIDL104)

Rankits R_i or `Normal Order Statistics` are estimators of standard Normal deviates (Fisher 1938 p 25, Teichroew 1956, Harter 1961 and 1970). The algorithm for calculating exact rankit values is rather laborious (Royston 1982b). HYGINIST uses 'read data statements' for MW20 (from Teichroew 1956 page 416 rounded at the fifth decimal) and for sample size of 21 through 2000 the polynomials of Royston (1982b, NSCOR2) are used.
see example 49

Contents - Index

Figure 22 Lognormal probability plot on 24 TWA8 hours methyl metacrylate (LEIDL104)



Contents - Index

Example 49 Making rankits by Monte Carlo simulation

Rankits can be made with a Monte Carlo simulation:

- generate a sample of M random standard Normal deviates Z_i (Abramowitz 1971, page 953, direct method), and arrange them in ascending order,
- repeat this procedure at least a thousand times,
- the arithmetic mean of the 1000 deviates with rank i is an effective estimate of rankit R_i.

Contents - Index

B.1.2 Shapiro's W-test (uncensored)

"The one major drawback of the probability plotting technique .... is the lack of objectivity" (Shapiro 1990 page 21). Statistical tests on shape are more objective, because of their reproducibility. However, the classical tests on skewness and kurtosis (Fisher 1938, Geary 1947) are only useful for samples of at least 50 measurements (see Hawkins 1991 page 58). These tests are now overshadowed by the omnibus test (= combined test on kurtosis and skewness) of Shapiro (1965).

The W-test of Shapiro (1965) estimates the effectiveness of the linear regression through the rankits (Schneider 1986 page 182). This special case of the analysis of variance test uses the following algorithm:

$$W = \frac{\left[\sum_{i=1}^{M/2} S_i * \left(x_{M-i+1} - x_i\right)\right]^2}{\sum_{i=1}^{M} x_i^2 - \frac{\left(\sum_{i=1}^{M} x_i\right)^2}{M}} \qquad \text{B-1}$$

The coefficients S_i in formula B-1 are established using:

- S_1=1/sqrt(2) for M=3,
- Royston's polynomials (1995, AS R94) for M>4 and M≤2000.

The one sided probability distribution A(W) is (Royston 1995, AS R94) is calculated using:

- the exact W-distribution for M=3,
- a transformation to a standard Normal approximation for 3<M≤2000.

The power of the W-test is strongest against short-tailed (platykurtic) and skew distributions and weakest against symmetric moderately long tailed (leptokurtic) distributions (Royston 1992).
The significance level of A(W)<5% can be used to reject the Lognormal null hypothesis for MW 30. The effectiveness of the W-test decreases if results are rounded in grouping intervals ΔC (Pearson 1977). The W-test needs to be adapted (Royston 1989) if ΔC>0.1*ω (with ω= the arithmetic standard deviation see formula B-18).

Contents - Index

---

note 28

formula B-1 shows that the W test statistic is the ratio of two different estimators of the variance in a sample. The F-ratio test is based on the ratio of the variances of two different samples.

Contents - Index

Efficiency

According to Fisher (1921) an estimator is efficient if his population variance is the smallest (Kendall)

Contents - Index

Error

See effectivity

Error of the first kind

Rejecting a valid hypothesis. Also sensitivity (detecting non-compliance)

Error of the second kind

Accepting an invalid hypothesis. Also specificity (rejecting compliance)

Estimate

The value of the estimator (Kendall)

Contents - Index

Estimator

- A rule or method to estimate a constant in the population (Kendall)

2.2.4.2 Example files *.HYG

The program contains 36 files with industrial hygiene or quality control example data. More information on the example data files in Table 1. The name of the data files is mostly compiled from the authors name and the page number of the reference. The files in Table 1 are used throughout the text as examples. The descriptive statistics calculated for BOLEY62.HYG (see Example 57) and CHIP123.HYG will differ from those presented by the authors.

Table 1 Example data from different sources

Expectation

The arithmetic mean of a probability distribution

Contents - Index

A. EXPOSURE ASSESSMENT

HYGINIST is an instrument for the quantitative exposure assessment:

- Sample exposure data and the Lognormal distribution are used to draw general conclusions on exposure range, compliance with hygiene limits and inference with reference exposure characteristics.

This qualitative annex describes four aspects of exposure assessment:

- its place within the process "effective control of occupational exposure risks" (A.1),
- domains in which it can be utilized (A.2),
- different exposure assessment strategies (A.3)
- the extrapolation of sample data (A.4).

In Annex B the quantitative aspects of exposure assessment are discussed

Contents - Index

A.1 Effective control of exposure

The evaluation of exposure is part of the effective control of occupational exposure risk factors (working conditions). This is in many countries (IOHA Conference 1992) a part of Health & Safety control regulation and legislation. Effective working condition control can be schematized as a nine step process.

Table 13 Effective control of occupational exposure risks

```
+----------+   +------------+
|         1|   | Limit     2|
| hygienic +-->| Value      +--+  +--------------+
| hazards  |   +------------+  +->|  Exposure   4|yes/unknown
+--|-------+   +------------+  |  |  compliance  +---+
   |   |       | Exposure  3+--+  +------|-------+   |
   |   |       +-----|------+            |           |   +-------------+     +------------+
   |   |             |         +---------+no         |   |            8| yes | Effective 9|
   |   |       +-----+------+  |  +--------------+   +-->| Sustainment +---->| Control    |
   |   |       |           5|<-+  |mental &     7+---|   |             |     |            |
   |   |       | Control    +---->|physical      |   |   +-------------+     +------------+
   |   +------>| measures   |     |resistance    |   |          |no               no |
   |           |            |     +--------------+   |          |                    |
   |           |            |     +--------------+   |          |                    |
   |           |            +---->|Co-exposure  6+---+          |                    |
   |           +-----|------+     +--------------+              |                    |
   |                 |                                          |                    |
   +-----------------+----------------------<-------------------+                    |
   +------------------------------------------<--------------------------------------+
```

1 Hazards. The static, innate health and safety aspects of an occupational agent (physical, chemical, biological and mental).
2 Limit value. The intensity of an agent over a reference period that is legally, health-based or otherwise accepted as admissible.
3 Exposure. The, at large, conductable and dynamic intensity of the occupational agent in the workplace, caused by the plant, the organization and the conduct of the employees.
4 Testing the exposure against a limit value. This leads to the decision:

- no, additional control measures are necessary,
- yes, exposure is in compliance with the limit,
- unknown, the limit is not suitable to test the exposure (co-exposure, restricted resistance).

5 Control measures. The technical, organizational and restriction of emission or immission.
6 Co-exposure. The controllable, but static interaction with other occupational agents.
7 Mental and physical resistance of an individual or a group. The result of:

- innate health, unequally distributed in the population,
- homeostasis, the tendency to maintain the internal stability in a physiological system,
- health perception, the dynamic vision of a person on its health and/or on public health, and
- health risks perception, the dynamic vision of a person or a community of hygiene hazards

possibly threatening health.
8 Sustainment. Testing total exposure in similar exposure groups against the physical and mental resistance by means of health examination or epidemiological analysis. This leads to the decision:

- yes, there is a state of effective control,
- no, supplementary control measures are necessary.

In both cases this may increase the knowledge about hygiene hazards.
9 Effective control. A state of control in which:

- employer and employees are cognizant with and give sufficient credit to industrial hygiene risks and control measures,
- exposure is in compliance with the limit values,
- technical and measurements are in compliance with the current industry state of the art and comparable with what's usual practise in this branch of industry,
- occupation related complaints, effects and absenteeism and the industrial hygiene risk complies with the socially accepted level,
- knowledge is generated about how to control occupational exposure risks effectively.

Contents - Index

note 24

The nine process steps can also be applied in the effective control of environmental hygiene, process safety and quality.

Contents - Index

note 25

Combined exposure of several agents (WHO 1981, Scheffers 1985),

- exposure through different routes of entry (lung, mucous membranes, skin and/or intestines) (Manz 1987, Thomas 1986),
- specific sensitivity (ACGIH 1982, Stresemann 1988),
- novel work schedules (Roach 1977, Hickey 1977, Brief 1986, Eide 1990),
- physical exercise (WHO 1981),
- extreme climate (Hertig 1975, WHO 1981 p29),
- use of personal protection.

The WHO report (1981) contains a comprehensive reference list with inaccessible references from east European periodicals.

A.2 Domains utilizing exposure assessment

The evaluation of exposure in workplace atmosphere is practiced in Process safety, Industrial hygiene, health surveillance, Occupational medicine and Occupational epidemiology. In table 14, the application of exposure assessment is sorted by sampling plan and domain (Corn 1985, CEN 1995).

In the red colored fields of table 14 HYGINIST is useful for:
- the assessment of homogeneity in sample data (see chapter 3),
- the extrapolation to general results (see chapter 5),
- inference between two samples (see chapter 6),
- assessment of minimum sample size (see chapter 7).

The classification of exposure for a $TWA_{8\ hour}$ standard (column 2 of table 14) is explained in detail by Leidel (1977) and many others and is therefore not included in the program.

Contents - Index

A.2.1 Process safety

To control the state of a production installation (i.e. reactors, appendages, seals, gaskets, extrusions, sampling points, logistics) and to detect unwanted emissions, the chemical industry utilizes routine measurement programs on fixed places and with a fixed sequences, mostly with directly readable instruments. The grab samples are directly compared to the limit value. For there is a legal framework for continuous or sequential sampling (EG directive 78/610/EEG, PB L 197/12 of July 22, 1978).

Contents - Index

A.2.2 Industrial/Occupational hygiene

Industrial (US) or occupational (GB) hygiene, in its classic form, deals with the working conditions at the workplace: the spot where employees meet hardware and materials. The aim is to keep risks in compliance with the current standards in order to guarantee workers a safe and healthy stay in the work environment not only during a (fraction of the) work shift but even during a lifetime. The method is to recognize, evaluate and control the agents at the workplace.

Industrial hygiene is a mature applied science: You can find basic or complete professional information of almost every present-day problem in the peer review or other literature that is produced since the forties and that can be found in hardcopy, on-line or compact disk databanks. The recognition of chemical agents risk is described by e.g. Roach (1977), Leidel (1977 chapter 2), Clayton (1978 p29-112), Oostendorp (1985ab), Boleij (1987), Guest (1993), Hawkins (1991), Mulhausen (1998) . The statistical aspects of the evaluation of workplace exposure can be found throughout the complete text.

Contents - Index

A.2.3 Health surveillance in similar exposure groups

In chemical process industry the evaluation of exposure is often conducted among similar exposure groups (Mulhausen 1998, Hawkins 1991 page 5, Damiano 1987, Corn 1979). Its aim is:

- assessment of long-term risk of overexposure,
- administration of exposure history.

In occupational settings the activity pattern of groups of employees is often similar throughout the year, despite the sometimes significant between-worker-variability during short observation periods (Rappaport 1993). Within a SEG the distribution of activities among the workers throughout the year should only be determined by chance. The statistical aspects of assessing exposure and risk in similar exposure groups can be found throughout the complete text.

HYGINIST strongly supports the quantitative aspects of SEG health surveillance programs as described in the appendices of AIHA's strategy for occupational exposure assessment (Hawkins 1991):
- II, descriptive statistics and plotting probability,
- III, arithmetic means tests,
- IV, tolerance limit test.

A.2.4 Occupational medicine

Exposure data are sometimes used in assessing the sustainment (see table 13) of a healthy individual or in employees with a limited mental and/or physical resistance against the total of occupational exposure. According to Rempel (1990 page 435): "The principal tool for preventing work related illness should be the industrial hygiene exposure assessment (primary prevention). Unfortunately, even in the best situations, work related illness occurs. Medical surveillance in the workplace functions as a backup to exposure assessment, as a 'safety net' to catch illness early and activate interventions to prevent it from progressing (secondary prevention)." Statistical evaluation of exposure data will be of limited use in this domain.

Contents - Index

A.2.5 Occupational epidemiology

Occupational epidemiology (Monson 1980, Smith 1987, Checkoway 1989, Harris 1993) applies exposure assessment in similar exposure groups, to find:

- dose-response relations between exposure and health (biometrics, blood and urine values, lung function, mortality and morbidity rates),
- no-effect or accepted risk levels for industrial hygiene limit value assessment.

Table 15 Scheme for health based limit value setting in similar exposure groups

Contents - Index

Table 15 Scheme for health based limit value setting in similar exposure groups

```
 +--------------+
  ¦Complaints or ¦
  ¦Medical occu- +---+   +---------------+     +--------------+
  ¦rences        ¦   ¦   ¦ Dose/Response ¦     ¦ Occupational ¦
  +--------------+   +--?¦ relation      +----?¦   Limit      ¦
  +--------------+   ¦   ¦               ¦     ¦  Value       ¦
  ¦Exposure      +---+   +---------------+     +--------------+
  ¦assessment    ¦
  +--------------+
```

Contents - Index

Table 14 Exposure assessment applications by domain

Sampling plan One grab or TWA One series D2 series exposure situa- com-
meassurement. exposure data. tions. Comparing differen- ment
ces in regimes of control
Domain
Process Incident, accident Continuous and Technical measures A.2.1
safety/emission and leakage semi-permanent
Control monitoring
Systems

Industrial hygiene Compliance with a Long term Technical and. A.2.2
(workplace oriented) Standard compliance control Organisational
Control charts

Health surveillance Range finding General Technical, organisational A.2.3
Similar exposure assessment and and behaviour measures.
Group compliance control
Occupational Individual long term Sustainment under A.2.4
Medicine compliance occupational load Different circumstances
(single worker)

Occupational Range finding dose estimate Dose groups A.2.5
epidemiology and
industrial toxicology

Contents - Index

A.3 Existing exposure assessment strategies

In recent decades strategies have been developed for the long-term exposure assessment (e.g. AIHA, BIN, CEFIC, CONCAWE, EH42, Leidel 1977, TrgA 402, VDI 2450, WGD, WHO, CEN 689, Guest). In general these strategies all include four steps:

- what are the hazards (agents, routes of entry, exposure duration and frequency),
- what is the exposure (identifying similar place/group/time situations),
- what are worst-cases (ranking hazard and exposure),
- which of the ranked worst cases is in non-compliance with the limit value(s) ?

They all provide some sort of evaluation to generalize measurements performed.

In the last few decades the industrial hygiene world is flooded with alternative exposure assessment sampling strategies and statistical methods to analyze sample data. Some of these methods provide only concepts and do not even claim to improve the power of exposure control over the classical methods with fundamental basic concepts. They are ignored here.

A.4 Extrapolation

The methods to generalize the measurement results in the assessment strategies are based on:

- the experience of "experts" (CEFIC, CONCAWE, TRgA 402, WGD),
- distribution free statistical analysis (VDI 2450),
- the Lognormal distribution (AIHA, BIN, CEN 689, EH42, NIOSH, WHO)
- other parametric distributions (Berry 1973, Owen 1980).

HYGINIST is based on the Lognormal distribution. The statistical aspects are explained throughout the manual. Other parametric distributions are explained only superficially (see B.4.2). Stellingwerf commented (1984, letter to the editor) the power and other statistical aspects of the first two methods.

Depending on sampling duration and limit value reference period, two types of extrapolation to unsampled periods (general results) can be distinguished:

- In number: the measured periods are considered as a fraction of the total number of exposure periods,
- In time: the sampling durations are considered as a fraction of the total exposure reference period (see Example 47).

For extrapolation in numbers the sampling duration should be equal to the limit value's reference period.

Contents - Index

A.4.1 The Lognormal evaluation

Most assessment strategies are based on Lognormally distributed workplace air concentrations (among others BIN, Bar-Shalom, EH42, Leidel, Rappaport, Tuggle, WHO, CEN 689).

According to Esmen (1977) and many others population, mutually independent, industrial hygiene measurements are best represented by the Lognormal distribution (see Figure 21). `Best` means, according to Esmen (1977), that the estimator is unbiased or consistent and efficient. `Best` also means that the estimator has the highest probability among the existing estimators. (Gamma, Weibull, Normal, exponential, Nonparametric).

Much knowledge and experience on the Lognormal distribution, extrapolation and inference has been gathered within the biometrics (Gaddum 1945), econometrics, technometrics, survival analysis, meteorology and geology (Shimizu 1988). Industrial hygiene has learned and is still learning from these applied sciences.

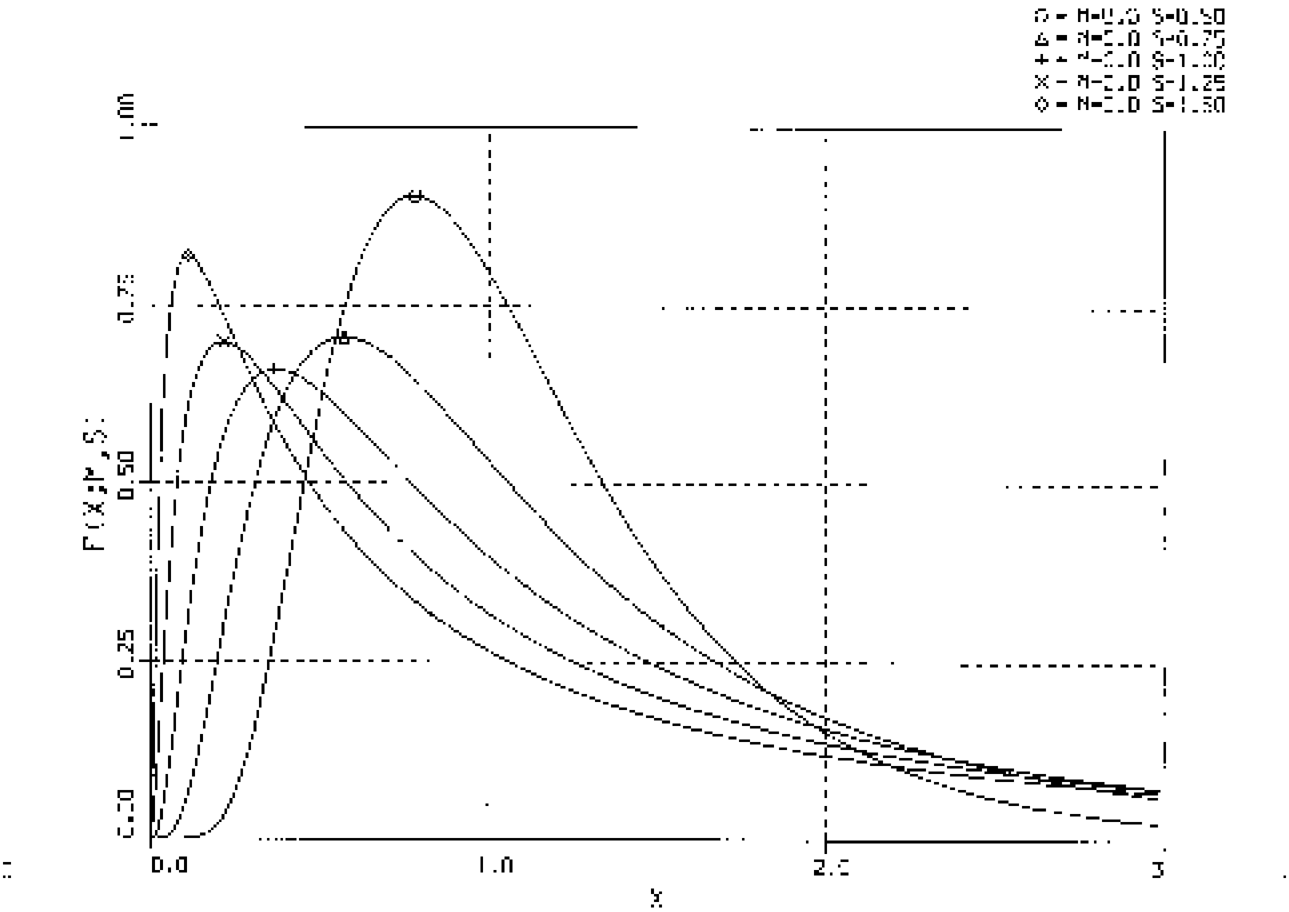


Figure 21 A family of Lognormal distributions with mean M and standard deviation s

[Example 48 Lognormal]

Contents - Index

Example 48 Lognormal

The Lognormal distribution in series industrial hygiene measurements was first recognised in the fifties at the coal face (Oldham, 1953). Lognormal goodness-of-fit has been demonstrated in the distribution of concentrations of gasses, vapours, dust, fibres and enzyme-activity in workplace air, in excretion products in urine, in exhaled air, in blood values and in (untransformed) noise intensities (kPa). Leidel (1975b page 7, 1977 page 122) indicates that Lognormal goodness-of-fit is observed in American, English and German industry. The Lognormal distribution is observed in series short-term (seconds) and long-term (days) sampling periods (Leidel 1975, 1977).

Contents - Index

A.4.2 Other parametric distributions

The Lognormal distribution Λ can be approximated by a linear combination of Normal distributions (Shimizu 1988a page 9):

$$\Lambda = \eta + \frac{(\eta^2 - 1) * \sigma}{2} + \frac{(2 * \eta^3 - 9 * \eta) * \sigma^2}{12} + O(\sigma^3)$$

with Λ the Lognormal L(μ,σ) and η the Normal *N*(μ,σ) distribution. If $\sigma<0.2\rightarrow 0$, then the last three terms tend to zero and the Lognormal distribution is approximately normally distributed.

Berry (1973, cited by Leidel 1977 page 103) finds, for certain contaminants in food and drinking water, the Gamma distribution more suitable than the Lognormal. He further expands this for samples with a high proportion of censored data. This method is not further applied in industrial hygiene.

The three parameters, Lognormal Delta distribution is effective in a stationary process with a constant low level background concentration (Owen 1980, Leidel 1977 104). In contrast to the statistical point of view, the three parameter Lognormal distribution is in industrial hygiene a special case of the two parameter distribution (with δ=0) because a random and a fixed exposure level are considered.

According to Shapiro (1991 p8, 15) many of the distributions applied are mutually dependent:
- the exponential distribution is a special case of both the Gamma (Chi-square with 2 degrees of freedom) and the Weibull distribution,
- the Weibull becomes an "extreme value distribution" through logarithmic transformation,
- the Lognormal distribution can be approximated with a Weibull or an exponential distribution.

In conclusion there is no evidence and there are no good reasons to use other distributions in stead of the Lognormal in the industrial hygiene exposure assessment.

Example 47 Extrapolation in time

The reference period of the following limit values obliges you to perform (1) a variable sampling plan and (2) extrapolation:
- the Dutch limit value for lead TWA40 hour,
- the EG asbestos cumulative action level for 3 months,
- the EG vinyl chloride limit for one year.

The Dutch MAC30 min for carbon monoxide is often measured by a series grab IR registrations. Using IR or indicator tubes makes it usually impossible to sample the total reference period ($\sigma ti << T$), so extrapolation in time is preferred. Statistical techniques are available to evaluate the TWA from partial-period or grab sampling (see 5.2).

B THE LOGNORMAL DISTRIBUTION

The statistical evaluation of industrial hygiene sample exposure data is tied to different, implicit and explicit, conventions. Which convention or alternative HYGINIST uses, is made explicit in table 16 and is explained in the chapters mentioned in the last column.

With good Statistics , population characteristics like Goodness-of-fit, location, range, and inference can be estimated from a sample, in a reliable and reproducible way. It is the task of the industrial hygienist to develop a reliable exposure assessment strategy and to choose the right statistical methods. Therefore, good exposure assessment requires both industrial hygiene and statistical skills.

Proposition 1:

- **Quantifying workplace air exposure in similar exposure groups is most effective with a random or stratified sampling strategy and with Lognormal extrapolation.**

HYGINIST evaluates the following situations:
- A complete sample to a complete population (B.2),
- A censored sample to a complete population (B.3.),
- A complete or censored sample to a confined population (B.4.2.).

Proposition 2:
- **Lognormal extrapolation is possible if the expert industrial hygienist decides that the model is acceptable., based on:**
- **the population description,**
- **the measurement strategy,**
- **the effectivity of the measurement method, and**
- **the goodness-of-fit,**

This proposition makes a heavy appeal on the experience and expertise of the industrial hygienist. A scheme which helps to decide on conformity with Lognormality based on statistical grounds is presented in par. 3.4 . In non-experimental research, like industrial hygiene and occupational epidemiology, elements are not randomised over the populations (Miettinen, 1985) and thus the variance in observations on elements can be in part non stochastic. The bias in population selection, exposure assessment strategy and the measurement method will distort the conclusions. Checking on selection is more than just goodness-of-fit control: it means auditing the complete design of the investigation.

Proposition 3:
- **Unbiased estimators and unbiased extrapolation methods should be preferred.**

This proposition deals with the industrial hygiene habit of general extrapolation from small samples (M<10). A perfect estimator is unbiased with minimum, and normally distributed, variance. The inaccuracy of maximum likelihood estimators in small samples leads to more errors of the first kind (incorrectly accepting compliance). This should be considered as more serious than the lower efficiency (with the chance of incorrectly accepting non-compliance) of unbiased estimators.

Proposition 4

- **If data are censored, estimating the descriptive statistics from the rankits plot is the method of preference.**

Eliminating censored data or using a fraction of the detection limit as their value is replaced by a more effective method (see B.3)

Proposition 5:

- **It is general practice in industrial hygiene to perform statistical extrapolation to general results.**

The extent of the extrapolation and associated degrees of freedom is further elaborated in B.4.

Proposition 6:

- **In censored samples the number of results outside the accuracy range contributes for one half to the number of degrees of freedom with a maximum equal to of the number of the uncensored results.**

This proposition is a compromise between extremes (see B.4.3).

Proposition 7:

- **Extrapolation to a one-sided 95% confidence level is just tradition.**

Almost every application of statistics uses 95% as default for inference and extrapolation (see 4.2).

Proposition 8:

- **Student-t based statistics should be preferred in industrial hygiene statistical inference and extrapolation.**

The Student distribution is asymptotically normal, sufficiently robust to guard against violating the normal model (Boneau 1960, Heeren 1987) and relatively easy to calculate.

Table 16 Conventions in industrial hygiene statistics

Characteristic Convention Alternative explanation
propo Pragraph
sition

The shape of the Lognormal Normal, Nonparametric 1 A.4
distribution Gamma, Weibull

Test for goodness (Log)normal (median describing population homo 2 B.1
normalorder) probability geneity, Rankit probability
-of-fit paper. isher and Geary tests plot, W-test.Waters test
onkurtosis and skewness

Descriptivestatis consistent, efficient Unbiased 6 B.2
tics

Coping with $C_{<dl}$ = ½ C_{dl} Maximum likelihood or 4 B.3
censored data dl= accuracy limit Rankit regression through
the uncensored data

Degrees of extrapolation to extrapolation to a 3 B.4
freedom df generalresults. confinedpopulation
df=M-1 M?:df=(M?-1)(M-1)/(M?-M)

df in censored all data df=M-1 or the the minimum value of 5 B.4.3
samples uncensored only: df=M\`-1 df=(M+M\`)\2-1 or df=2*M\`

Extrapolation Extent upper one sided 95% two sided, different percentages 7 4.2

Extrapolation standard Normal unbiased and 8 4
Technique Noncentral Student

B.3 Parametric descriptive statistics (censored sample)

Censoring (Hald 1949) is a situation in which a part of the sample data has an unknown or a fixed value. Type I corresponds with censoring with accuracy limits (Schneider 1986 page 2). Methods based on censored samples have the advantage of resistance to outliers (Snedecor 1980 page 135): A small percentage of outliers, produced by gross mistakes in measuring, recording and copying, will hardly bias the estimators.

See:

- B.3.1 on how Lognormal descriptive statistics are calculated using linear regression through the outcome within the detection range see
- B.3.2 on how the other descriptive statistics are calculated

Censoring greatly complicates extrapolation and inference (Schneider 1986 p 177). With asymmetric (one sided) censoring, a case that is often seen in industrial hygiene, the estimators of the descriptive statistics μ and σ become mutually dependent. The precision of censored based estimators and estimates depends on (Schneider 1986 4.7):

- the sample size,
- the symmetry of censoring.

Figure 25 500 Rivet heads (HALD151)

Contents - Index

B.3.3 GMu, AM, and w² (censored)

In a censored sample, the descriptive statistics GM_u , AM_u and w^2 are calculated using GM_g in B.3.1 , GSD_g in B.3.2 and the degrees of freedom df of B.4.3.

Figure 27 Mice survival time (GUPTA271, Schneider 1986 p 69 & 88)

Example 53 Performance of estimators of censored sample descriptive statistics

Contents - Index

B.3.1 Censored sample estimators for EXP(s)

In the case of a censored sample the program estimates σ from the linear regression through the uncensored data M' in the rankits plot. Log(GSD_g) is the regression slope (Prescott 1970, section 5):

$$GSD^g = EXP\left(\frac{\sum_{i=ll}^{i=ul} R_i * x_i - \frac{\sum_{i=ll}^{i=ul} R_i \sum_{i=ll}^{i=ul} x_i}{M'}}{\sum_{i=ll}^{i=ul} R_i^2 - \frac{\left(\sum_{i=ll}^{i=ul} R_i \right)^2}{M'}} \right)$$

In formula B-20 *ll* and *ul* are the lowest and highest uncensored data rank and *M'* is the number of results between the accuracy limits. The formulas B.20 and B.21 are identical with the linear, alternative estimators of Gupta (1952 5.2. formulas 31 through 35) and Sarhan (1962 p208 formulas 10C.2.1 and 10C.2.2). See Example 14, 21 and 44 on the rankit estimators for mean and variance of strontium in milk.

Example 52 Comparing untransformed linear estimators

For uncensored samples the rankit estimators are as effective as the classical estimators (see B.2). They are more effective than classical solution where results below the detection lower limit LL receive the value .5*LL (Hornung 1990, Hawkins 1991 p 104). Example 53 suggest that the rankits estimators should be preferred in industrial hygiene.

Contents - Index

Example 44 Regression analysis in exposure control chart

*.HYG file Description
VINCHL## The TWA8 hour Vinylchloride VCM in a SEG polymerisation shift workers since 1987 (see Example 29) were presented in an exposure control chart (Figure 15). In contrast with the comparison of the 1989 and 1991 data, the regression line delineate a significant decrease of exposure in time. The regression slope calculated using TRUE EPISTAT indicates that the geometric mean exposure between 1987 and 1993 decreased every year with 100*[1-exp(-.1961)]≈18%. (95% confidence bounds 12-24%).

Contents - Index

B.4.3 Censored

There is no convention in industrial hygiene on how results outside the accuracy range of (1) the sampling procedure and (2) the analytical method, should contribute to the degrees of freedom. The following methods are found (with no justification given for the choices made):

- Perkins (1990) establishes the arithmetic mean of a large, censored sample with the uncensored data only (df=M`-1).
- Schneider (1986 page 190) uses the uncensored data when sample size is small (df=M`-1).
- Schneider (1986 page 200 & 201) estimated tolerance limits (=percentile values), independent of sample size (df $\equiv \infty$) on 96 locomotives of which 59 were type I censored (large samples are considered as asymptotic normally distributed).
- Hornung (1990 method 2 & 3) and Hawkins (1991 p104) use both censored and uncensored data (df=M-1),
- The most liberal is the NIOSH method for non-compliance probability in long-term exposure situations (Leidel 1977 page 65) which is independent of sample size and accuracy range (df $\equiv \infty$),

If:
- sample data show enough variance (e.g. GSD>1.4)
- the standard deviation w>>-C overshadows the grouping interval, and
- limit H lies inside the accuracy range,

then censored data can be considered as data of equal value in a continuous distribution and thus, at least in part, mutually independent (see proposition 6):

$$df = \mathrm{INT}\left(\frac{M+M'}{2}\right)-1$$

However, if :

$$df > 2*M' \text{ then } df = 2*M$$

with M` the results within the accuracy range. This method is rather conservative taking into account that, for NIOHS` method of non-compliance probability in long-term exposure situations (Leidel 1977 page 65), no considerations on df have to be made (see screen 25 par.. 5.1.2). For Nonparametric (distribution free) methods dealing with censored data is more simple:
 - In the one sample binomial test and Poisson test with the hygienic limit inside the accuracy range, the results outside the accuracy range are considered as independent (df=M-1),
 - When the Wilcoxon rank test is applied to compare two samples and one sample contains nondetectable values, they are considered as equals and independent ($df_1=M_1-1$),

HYGINIST calculates the number of degrees of freedom for the censored and the non-censored case as follows:

```
  If DetecNumAantal < M Then
        Nu = (M + DetecNumAantal)  2 - 1
        If Nu > 2 * DetecNumAantal Then Nu = 2 * DetecNumAantal
    Else
        Nu = M - 1
    End If
```

With:
M is sample size
DetecNumAantal is the uncensored number of measurements
Nu is the Number of degrees of freedom

Example 55 Degrees of freedom in censored samples

Contents - Index

Example 55 Degrees of freedom in censored samples

- The 310 rivet heads with diameters (HALD151) above the lower limit of 13.4 mm (Hald 1952 page 151) in a sample of 500 (see Figure 24), provides a df=404.
- The 15 items under stress (SCHNE224) with an accelerated lifetime below the 6 time units (TU) in a sample of 50 (see Figure 5), provides a df=31 (Schneider 1988 p224).
- The 37 locomotives with failures before 135,000 Miles (SCHNE70.HYG) in a sample of 96 (see Figure 6), provides a df=65 (Schneider 1986 example 4.3 page 69).
- The 7 grab sample airborne Chlorine concentrations (OWEN716) within the range of accuracy from a sample of 15 that were taken over a working day (Owen 1980 page 716) in order to classify for the federal 8-hour TWA standard of 1 PPM (See Figure 4), provides a df=10.

Contents - Index

B.5 Sample size

In estimating exposure the following theoretical minima for the sample size exist:
- location, source strength, EXP(μ), β: one measurement,
- variance, scale, EXP(σ), w: two measurements,
- shape, kurtosis, goodness-of-fit: 3 measurements.

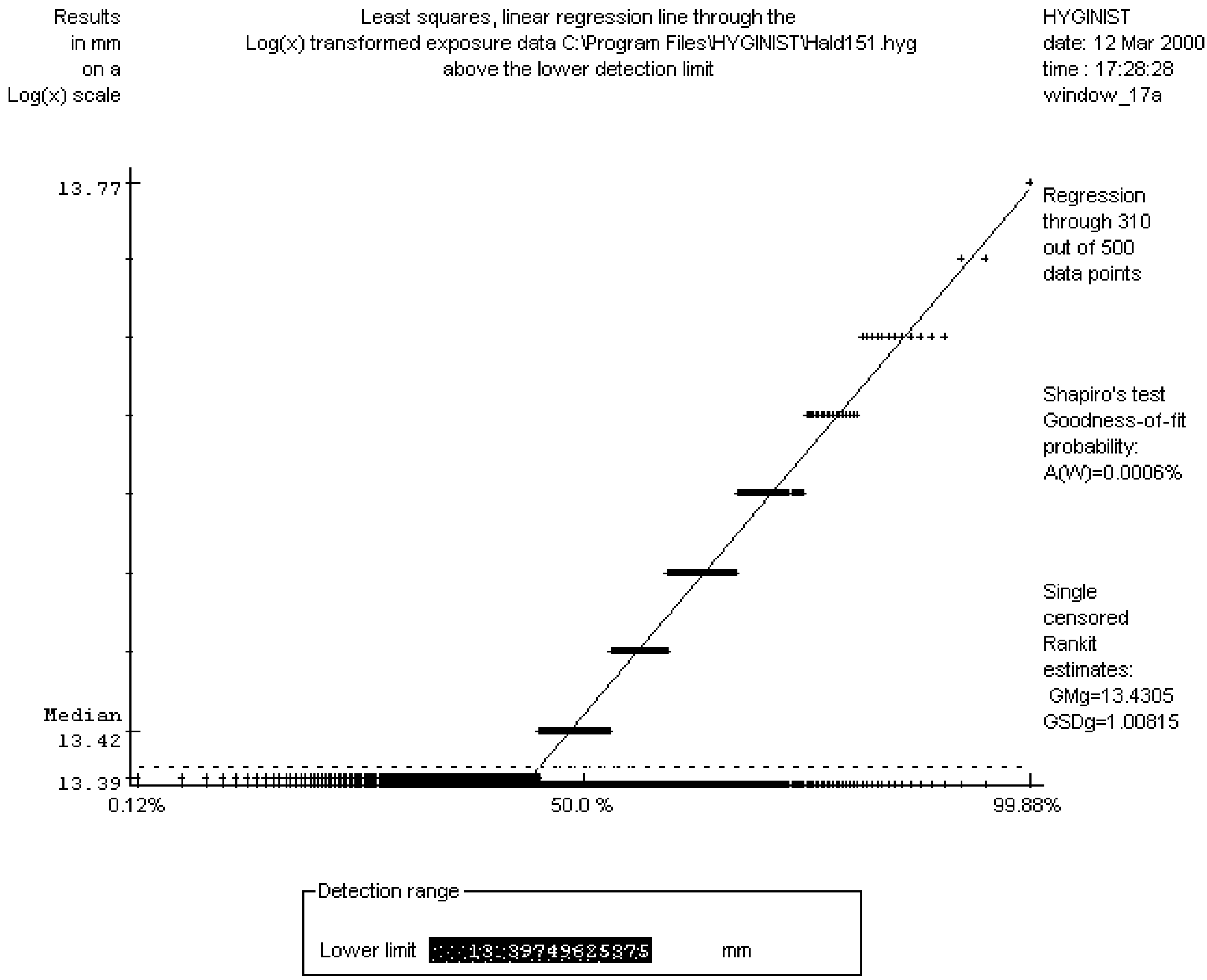
Most industrial hygienists will establish their sampling plan based on:
- the goal of the exposure assessment (see A.2),
- the measurement accuracy,
- the expected difference between concentration C and limit value H,
- organizational and financial considerations.

For every exposure application and domain (see A.2), sample size considerations are made in literature (Leidel 1977, page 120, The need for an action level; Leidel 1977, page 82, and Dewell 1989 accuracy based; Hawkins 1991, page 51, and Esmen 1992, long-term compliance control; Hawkins 1991, page 59; Schneider 1988, chapter 7, censored samples; Leidel 1977, pages 34 & 71, worst-case employee, cumulative binomial; Buhring 1989 and Hawkins 1991, page 59, variance based; Rappaport 1987, Am_{mo} dose based, criticised by Evans 1988; Corn 1985, page 176 and Tait 1992).

In general the sample size should increase if:
- the population under study is large,
- the desired confidence increases,
- the exposure variability is large,
- the difference between C and H is small,
- CV_t is large.

Contents - Index

Figure 25 500 Rivet heads (HALD151)



Contents - Index

B.4 Degrees of freedom

Degrees of freedom df are defined as the mutually independent variables of the sample in the population. If the statistical uncertainty (e.g. the extrapolation space) increases, the number of degrees of freedom decreases.
In a one dimensional sample df is established by:
- the number of measurements M (see B.4.1),
- the population size (expanded B.4.1 or confined B.4.2).

More complex are situations with:
- censored data (see B.4.3),
- when samples with unequal GSDs are compared (see 6.2.3.2).

HYGINIST calculates the number of degrees of freedom for the censored and the non-censored case as follows:

```
If DetecNumAantal < M Then
      Nu = (M + DetecNumAantal)  2 - 1
      If Nu > 2 * DetecNumAantal Then Nu = 2 * DetecNumAantal
  Else
      Nu = M - 1
  End If
```

With:
M is sample size
DetecNumAantal is the uncensored number of measurements
Nu is the Number of degrees of freedom

The choice of the estimators of the descriptive statistics and the extrapolation methods (see chapter 5) are, in part, determined by the size of df.

B.4.2 Confined population

If the reference period τ of a limit value (e.g. $TWA_{8\ hour}$, $TWA_{40\ hour}$ for lead, τ=480 hours or Mτ=60 shifts of $TWA_{8\ hour}$ in a period of 3 months for asbestos) can be divided in M? consecutive sampling periods (of e.g. t=1 minute, t=15 minutes or t=8 hours) and M non consecutive periods are sampled during time t (with M<<Mτ, Leidel 1977 p38), then the non-sampled periods over which extrapolation should be carried out is confined Mτ-M. The number of degrees of freedom in this case is (Juda 1967):

$$df = \mathrm{INT}\left[(M-1) * \left(1 + \frac{M-1}{M^{\tau} - M} \right) \right]$$

The last term in formula formula B-22 determines that df=M-1 for Mτ →∞ and df=∞; if M→Mτ. The influence of a confined population on df is relevant, if both Mτ and M are small; roughly if M<20 and MτW3*M.

Example 54 Degrees of freedom in confined populations

Contents - Index

Example 54 Degrees of freedom in confined populations

From Leidel (1977 page 63, 67):

- During 5 out of M_T=16 exposure periods to **Hydrogen sulphide** during one shift, a worker's exposure is sampled df=15*4/11≈5 (LEIDEL63.HYG).
- 10 PAS $TWA_{8\ hour}$ **Dioxane** in a period of 6 months (M_τ=130 working days). Confined as it is, the influence on the degrees of freedom is too small: df=INT(9.67)=9 (LEIDEL67.HYG).

40 $TWA_{8\ hour}$'s in a confined population of M_τ=200 shifts of one employee in one year results in df= 48.
5 $TWA_{4\ hour}$'s in a population of M_τ=13 Friday morning shift in one season results in df=6.

Contents - Index

6.2.3.2. Comparing GMs while GSDs differ

The statistic for testing that both GM`s originate from the same population base is (Snedecor 1980, 6.11):

$$t_{df_t} = \frac{LOG\left(\frac{GM_2}{GM}\right)}{\sqrt{(s+s_2)}}$$

with

$$df_t = INT\left[\frac{(s+s_2)^2 * df * df_2}{df * s^2 + df_2 * s_2^2}\right] \quad s = \frac{LOG(GSD)^2}{df+1} \quad s_2 = \frac{LOG(GSD_2)^2}{M_2}$$

Formule 6-5

Formula 6.5 follows, under the null hypothesis, the Student distribution. The two-sided probability A% is calculated using t_{df} and df_t from 6.5 and the Student distribution.

Example 41

Contents - Index

Example 41 Comparing GMs while GSDs are unequal

*.HYG file Description
VINCHL89& Because the exposure variance between 1991 and 1989 differs
VINCHL91 (see Example 39) the location parameters for $TWA_{8\ hour}$
Vinylchloride in the SEG PVC polymerisation shift workers were compared using the inference test of two GM`s with unequal GSD's. Although the ratio of GM_{1989}=0.83 PPM and GM_{1991}=0.37 PPM is 2.2, a significant difference is not confirmed: the two-sided probability for df_{pooled}=27 is $A_{GM=GM2}$=5.1%.
- Illustrating the influence of unequal GSD's in comparing GM.
For M=M2=10, GSD=2.71828 and GSD2=7.43714 differ significantly. $A_{GM=GM2}$W5% needs GM2/GMD4.6. This is 1.8 more than what is necessary if the GSD's are equal (see paragraph 6.2.3.1 and examples 40 and 41).

Contents - Index

Example 39 Comparing two GSD's

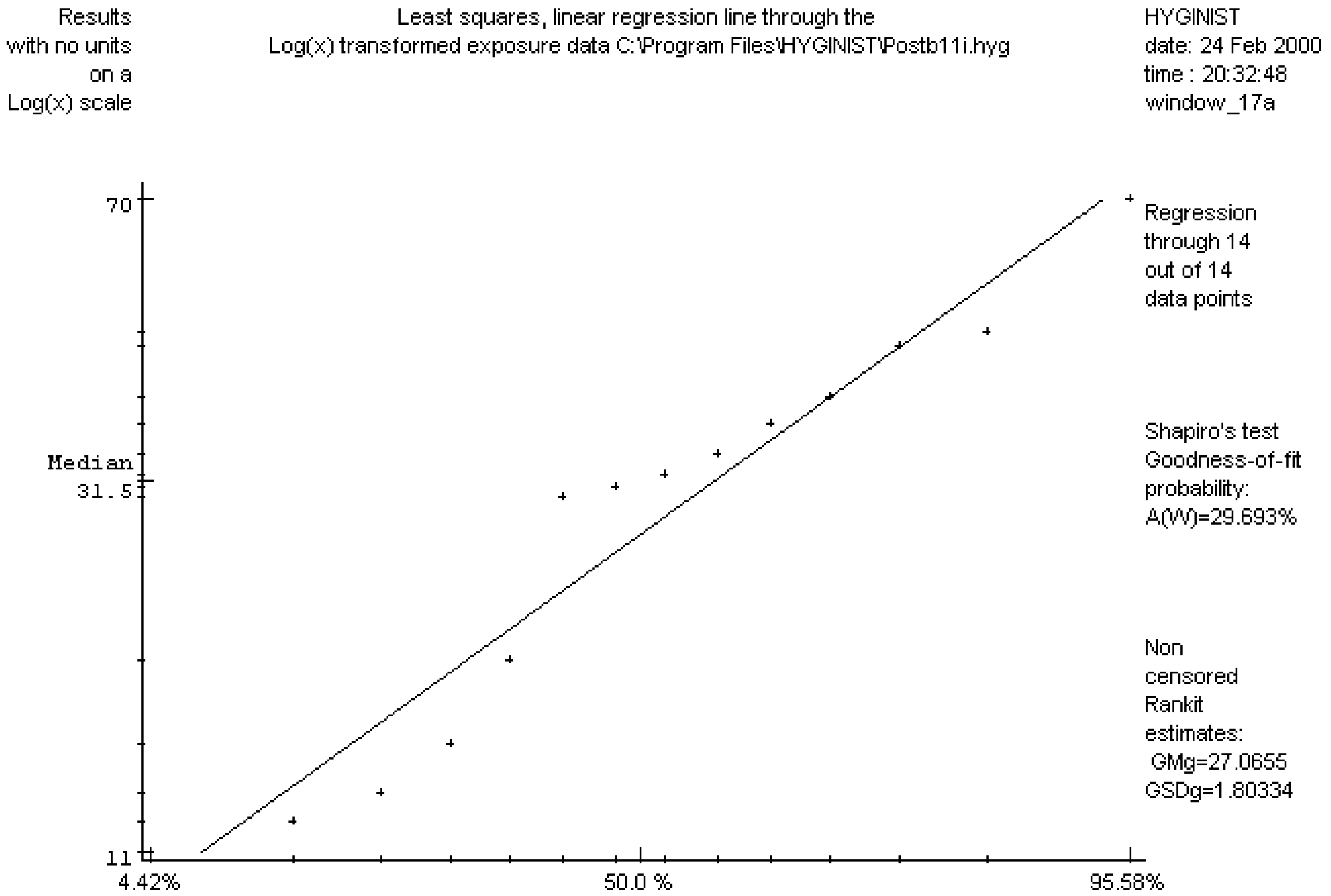
*.HYG file Description
POSTB11I, In a plant manufacturing fibreglass reinforced polyester materials,
POST08P4 14 Styrene PAS TWA8 hour concentrations were collected within a 15 days period
&POST10P5 on 7 inlayers working at 5 different presses (Post 1989, 1991).
The Lognormal probability plot (see Figure 3) suggest a mixture of at least two distributions and the raw data in the base report support this (Post 1989).
Four TWA8 hour concentrations from worker 8 at press 4 are compared with the five TWA8 hour of worker 10 on press 5 (measured on nearly the same days).
The scale parameters GSD=1.21 and GSD2=1.14 (see screen 29b) are from the same base: the two sided change AGSD=GSD=57.4% (continued in Example 40).

VINCHL89 & VINCHL91 The exposure control chart Figure 15 of the TWA8 hour Vinylchloride in the SEG PVC
polymerisation shift workers since 1987 shows a loglinear decrease in exposure over time. To quantify the decrease the 37 data of 1991 are compared with the 20 results of 1989. The scale parameters GSDdf=36=2.73 and GSDdf=20=4.97 differ significantly with a two-sided probability AGSD=GSD2=1.5 % (continued in Example 41).

- In the 1972-1974 randomized experiment in Colorado (Crow 1977 page 973), the mean hail mass on Mn=16 nonseeded days was GMn=2.3632 with GSDn=7.9135 and on the Ms=17 seeded days GMs=3.3434 with GSDs=8.35.
Like Crow (1977), the calculated AGSDn=GSDs=92.53% supports the conclusion that seeding does not influence the scale parameter of hail mass (continued in Example 40).
- From the Variance ratio distribution table 26.9 of Abramowitz (1970), it can be derived that two GSD's differ at a two-sided significant level of A=(2*QF¦df,df2)=5% for GSD=2.71828,
- df=df2=1 and GSD2=1.11*10+11,
- df=df2=8 and GSD2=8.21,
- df=df2=30 and GSD2=4.22.
To establish differences in exposure variability with 1.4WGSDW3.5, large samples (M>30) are needed.

Contents - Index

Figure 3 14 TWA8 hour Styrene on inlayers working on different presses (POSTB11I)



Contents - Index

Example 40 Comparing GM's while GSD's are equal

*.HYG file Description
POST08P4 & The Styrene PAS exposure level GM=38.6 PPM (see Example 9) from
POST10P5 worker 8 at press 4, is compared with GM=12 PPM of worker 10 on
press 5. Assuming equal GSD's (see Example 39) the two-sided
change $A_{GM=GM2}$=0.002% supports rejecting the null hypothesis:
The exposure level of inlayer 8 at press 4 is D3 times higher.
MOF134NS The effectivity of scavenging on the Methoxyflurane exposure in the
MOF134S breathing zone of veterinarian in small private practice, animal clinics
during surgeries was studied (Potts 1988 page 134). A comparison was
made between two series TWA10-30 min's measurements:
- 16 nonscavenged results (10 double censored data, Figure 19,
df=12) GM=0.48 PPM and GSD=5.7,
- 6 scavenged results (uncensored, Figure 20, df=5)
GM2=0.24 PPM and GSD2=2.9.The Wilcoxon rank sum test,
used by the authors, showed no statistical differences between
the two samples. Due to small sample size and large GSD the parametric
analogous with matching GSD's ($A_{GSD=GSD2}$=29.4 %) also supports the
null hypothesis: scavenging does not influence exposure level ($A_{GM=GM2}$=38.3 %).

- The median hail mass GM_n=2.36 on M_n=16 nonseeded days does not
differ from GM_s=3.34 on the M_s=17 seeded days in the hail seeding
experiment in Colorado, (Crow 1977 page 973, example 5, exponential
transformation). The two sided probability A(GMn=GMs)=63.8%, based on
equal GSD's (see Example 39), supports the null hypothesis.

- From the Student table 26.10 in Abramowitz (1970), it can be derived that
two GM's, from samples of equal size and scale, differ significantly at
two-sided A(t¦dft)=5% with GSD=2.71828,
- dft=2 and GM/GM2D 73.9
- dft=16 and GM/GM2D2.72
- dft=60 and GM/GM2D1.69.
With dfD10 and GSD≈2.7 significant differences in exposure level between
samples are easy to assess. See also examples 41 and 42

Example 42 Comparing two GSDs

*.HYG file Description
POST08P4&POST10P5 The GM and GSD's of Styrene PAS $TWA_{8\ hour}$s of worker 8 at press 4
and worker 10 on press 5 (see Example 40, Example 39 and Post 1989),
are combined. The summed sample size is M+M2=9, the merged
GM_t=23.5 PPM and the merged GSD_t=1.8. These values are exactly
the same as when all raw date are combined (POST1008.HYG)

- In a hypothetical fertilizer plant the plan is to perform a routine
program on total dust air sampling. In order to estimate an optimal
sample size two series of dust measurements from the past are
available, however, only the descriptive statistics were reported, the
raw data were lost. The sample estimators are:
- M_1=4, GM_1=1.8 mg/m_3 and GSD_1=2.4
- M_2=4, GM_2=2.24 mg/m_3 and GSD_2=2.4

Combining these descriptive statistics results in the following M_t=8,
GM_t=2.00798 mg/m_3 and GSD_t=2.26803. Note that with equal GSD's
and GM's which are very close, the combined GSD_t is smaller than in
the separate samples! See Example 46 for the assessment of the
optimal sample size.

Contents - Index

note 27

Statistics is a belief among experts. It is based on knowledge, logic and reproducible observations. Sampling is an art. Computing is a rigid, and in essence, simple skill. Fundamental sciences, and statistics in particular, are often misused. See e.g. the commonly found booklets "How to lie with ....".

Contents - Index

Extrapolation

- To estimate the value of a variable outside its tabulated or observed range

Contents - Index

Fraction

The number of elements of a population with a given property, divided by the size of the population

Contents - Index

Inference

The process of deriving from assumed premises either the strict logical conclusion or one that is to some degree probable (The Random House college dictionary)

Interpolation

To insert, estimate or find an intermediate term (in a sequence)

Contents - Index

Table 3 Type of distribution according to the goodness-of-fit with a transformation

Best transformation fit	Distribution type
C_i	Normal, Gaussian
$\log(C_i)$	If $s/x \approx$ constant for every C_i, then Lognormal
$\mathrm{Sqrt}(C_i)$	If $s^2 \approx x$ for every C_i, then the Poisson distribution
$1/C_i$	If $s \approx x^2$ for every C_i, then inverse normal

---

Kurtosis

The fourth moment

Contents - Index

Location

Mean (Sarhan), the first moment

Lognormal descriptive statistics

Location and variance parameters describing a Lognormal distribution

Minimum variance estimate

A method of calculating an estimator with the characteristic of maximum efficiency

Monomorphic group

A group of workers whose individual mean exposures compromise a single lognormal distribution. Term introduced by Rappaport (1991 page 66) but should be avoided. See Similar exposure group.

Contents - Index

Mutually independent

In opposite with autocorrelation

Contents - Index

non-compliance probability

The population fraction measurements over the industrial hygiene limit value

---

Occupational exposure assessment

Comparing exposure with the limit value (prEN 689)

Contents - Index

Omnibus

Term used by d'Agostino (1971) to characterise tests on shape that combine aspects of both skewness and kurtosis

---

Performance

General requirements on the effectiveness of the exposure assessment (prEN 482)

Periodic measurements

The regular check if exposure conditions have changed (prEN 689)

Contents - Index

Power

The chance of rejecting a unvalued null hypothesis

Contents - Index

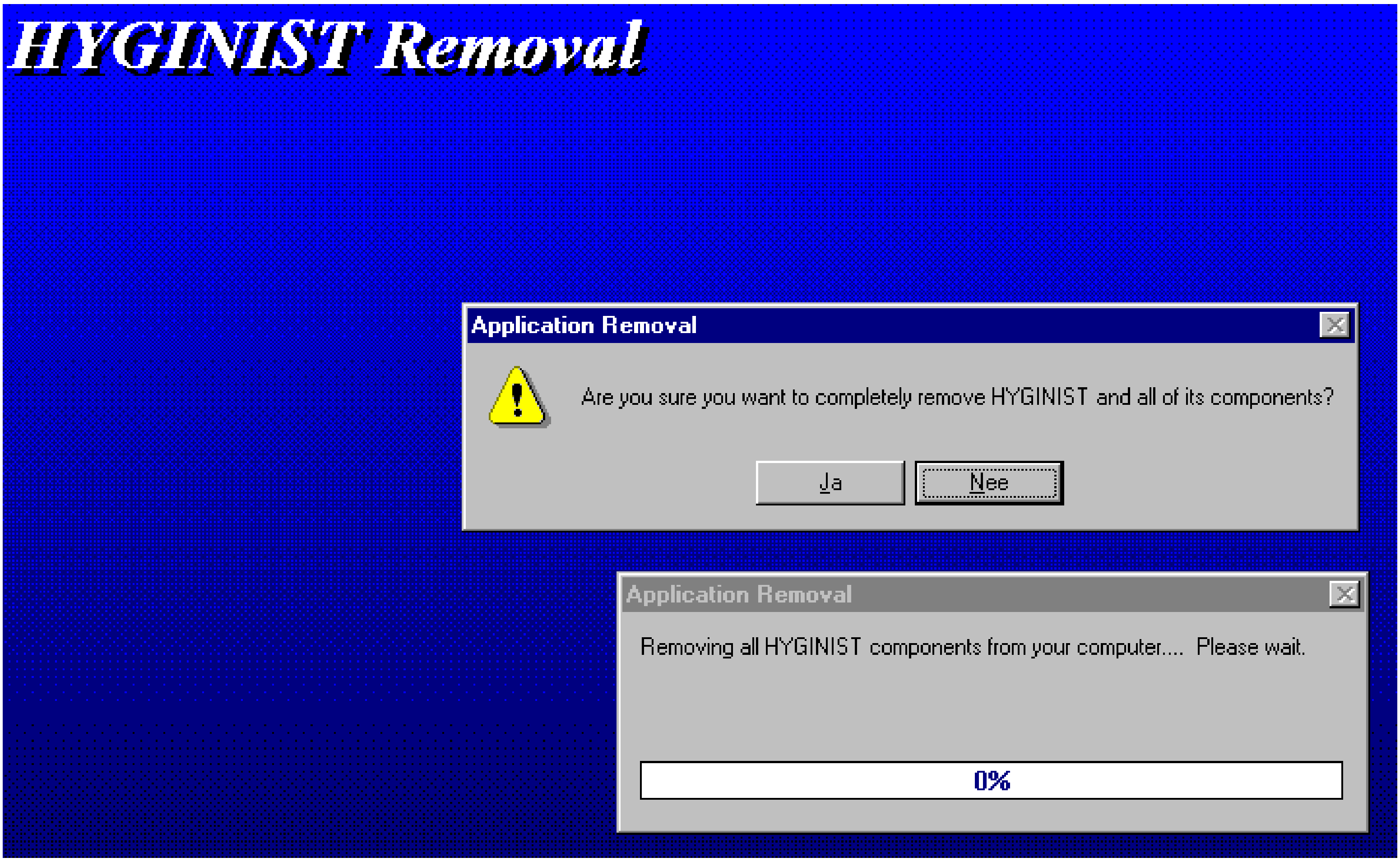
Precision

- Random error
- The size of the deviation from the mean of the observations (Hawkins 1991 163)
- The closeness of agreement between the results obtained by applying the method several times under prescribed conditions (prEN 482)

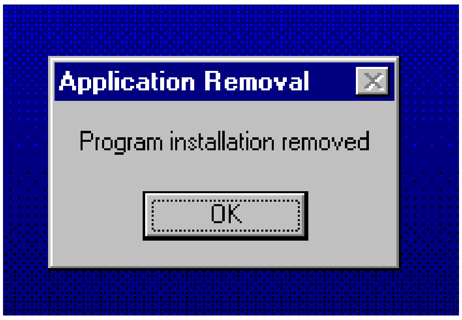
Uninstall or remove (earlier versions of) HYGINIST

If you have already installed a development version of HYGINIST for Windows or you want to remove the HYGINIST program from your system, then click in ascending order:

- Start,
- Preferences,
- Configuration,
- Software,
- Scroll to HYGINIST,
- Click Add/Remove.



All installed HYGINIST components are now removed.



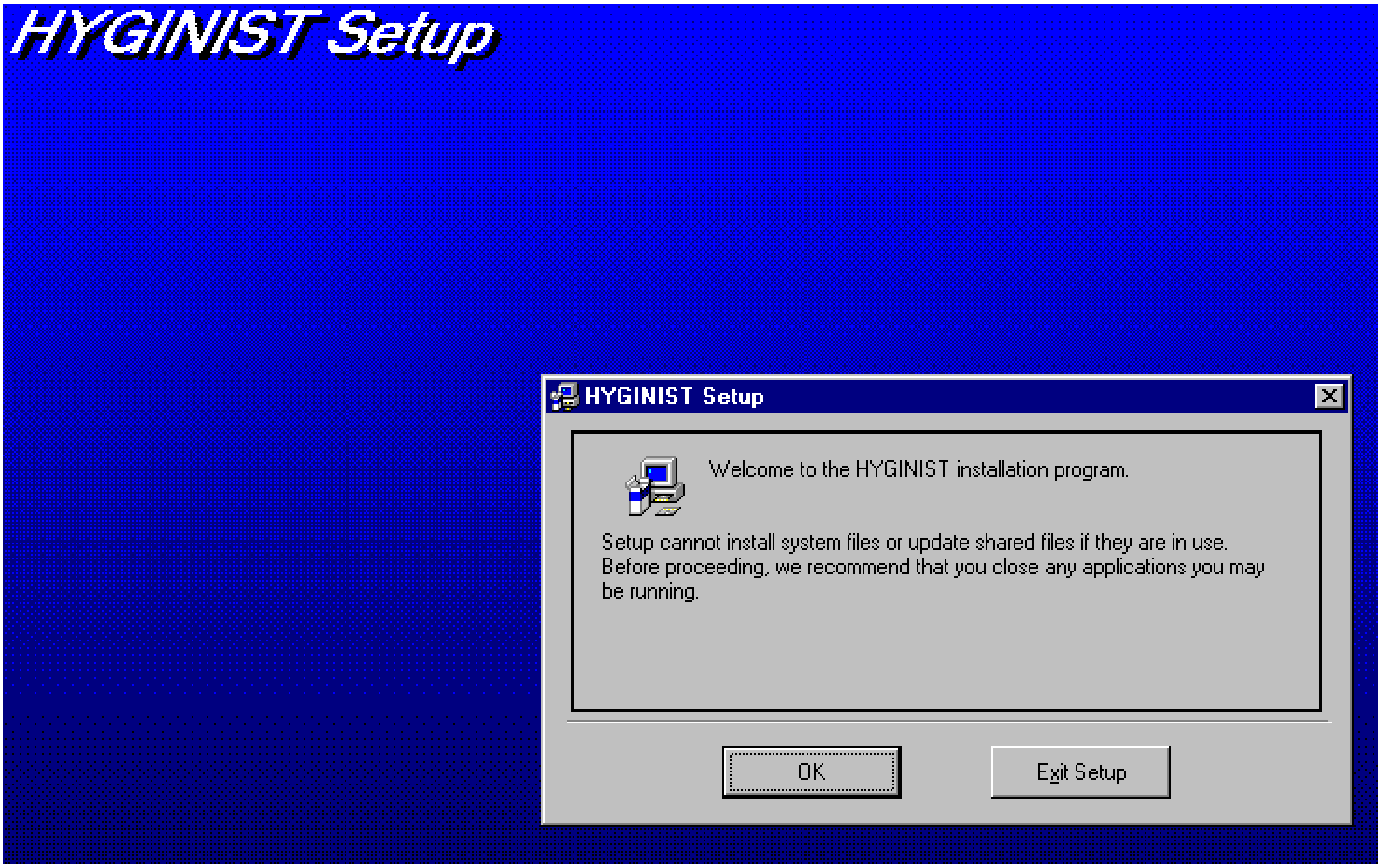
The remove is finished. Now continue with the installation.

C.4 Automatic installation on a hard disk

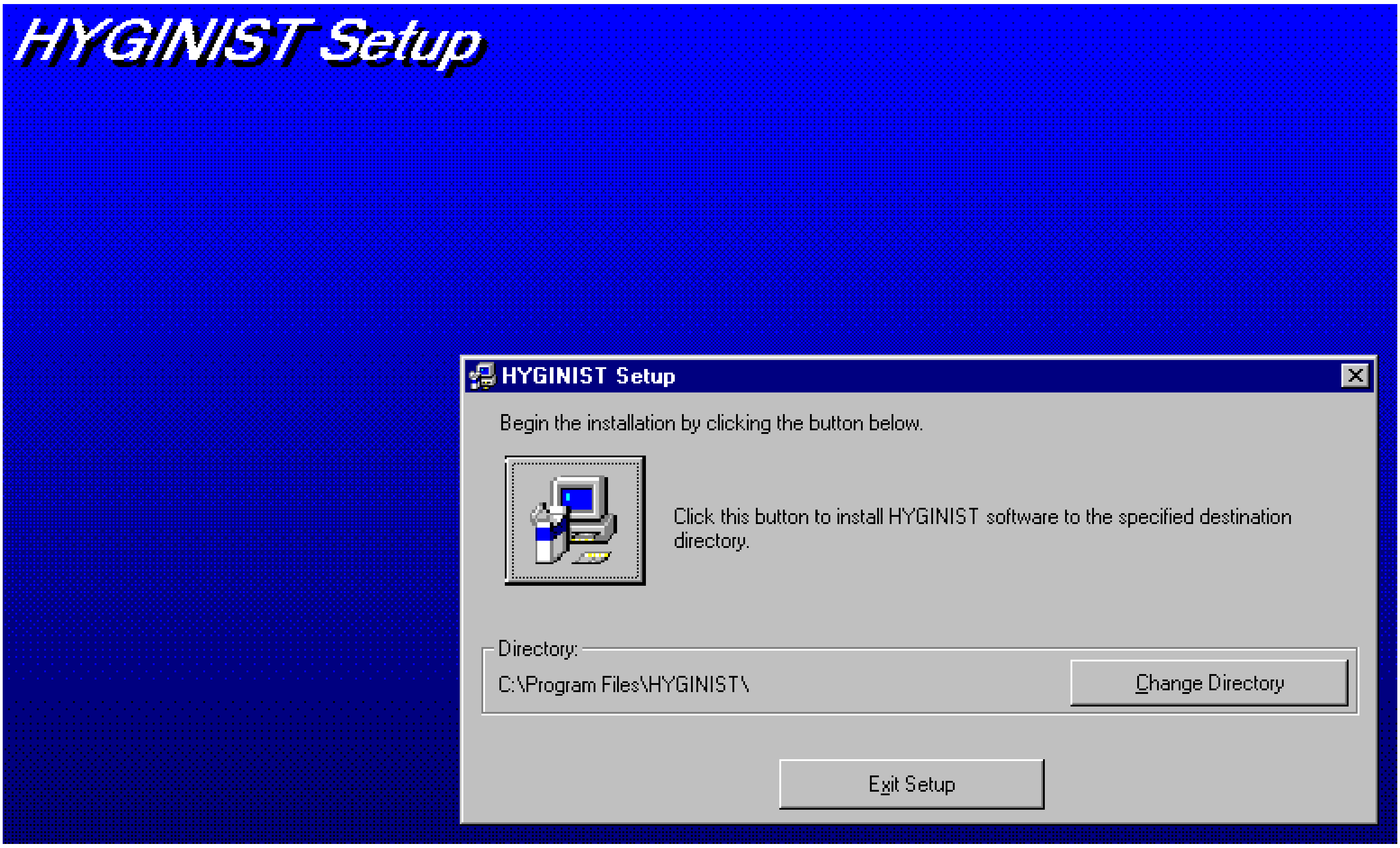
Browse with the explorer to the directory containing the extracted HYGINZIP.EXE files or:

- Click Start in the lower left corner and then click on Run,
- Browse to the directory containing the extracted files,

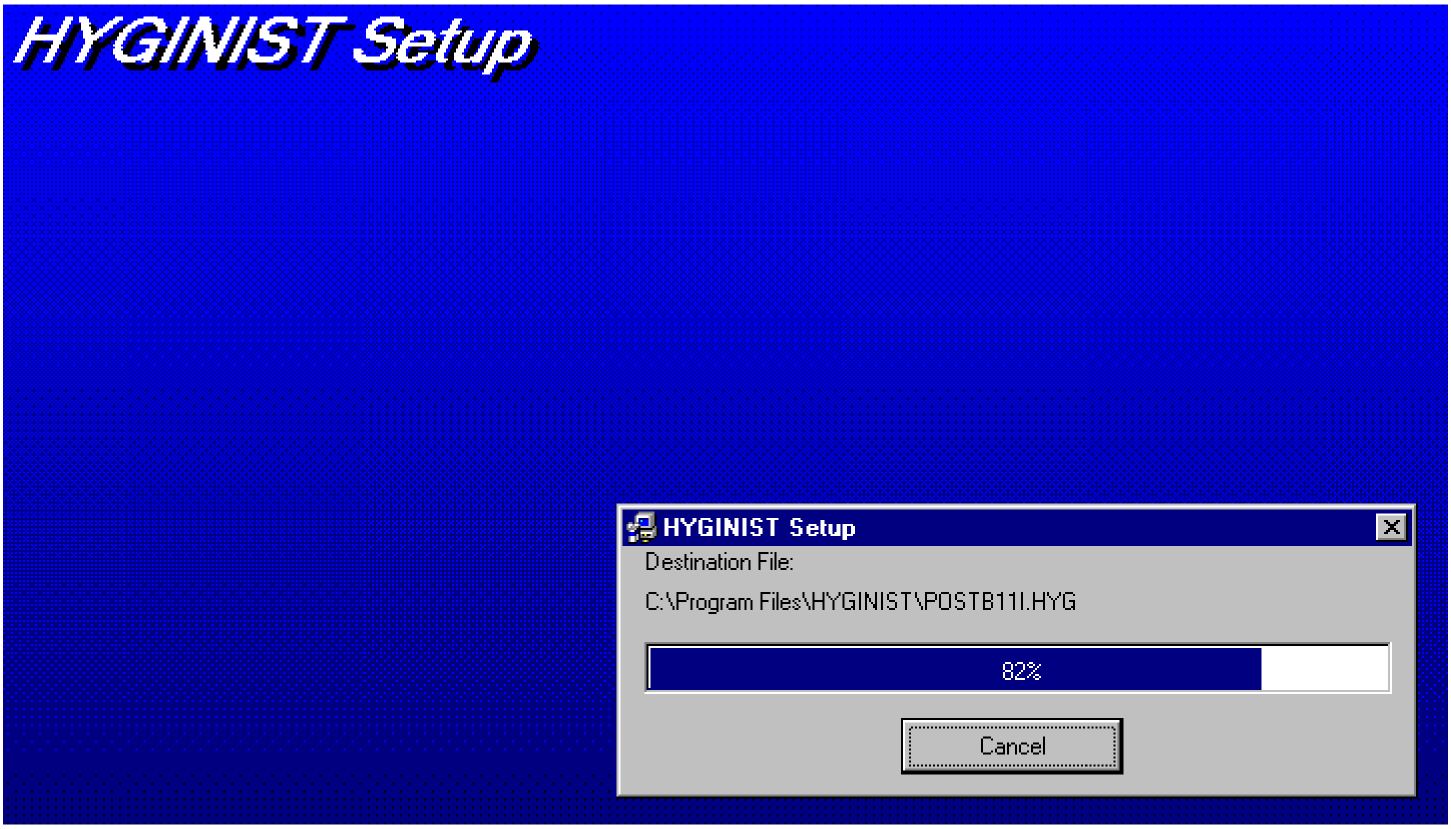
Now click on SETUP.EXE to start the installation.
The Installation starts with the following message window:



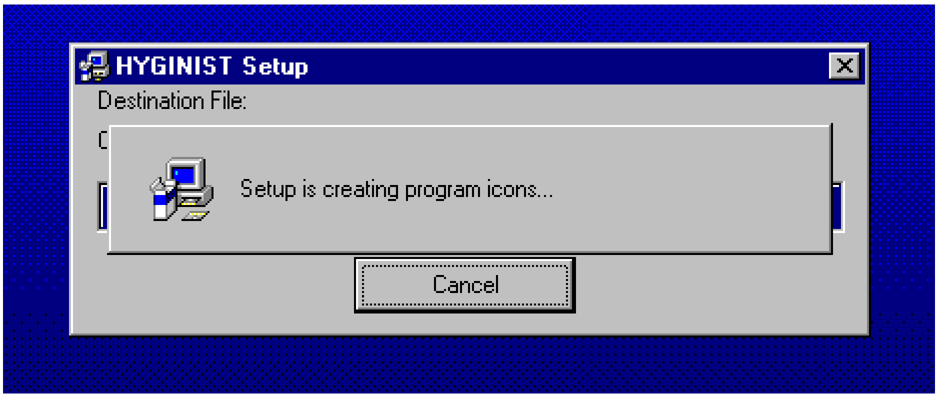
Close all programs before you continue.



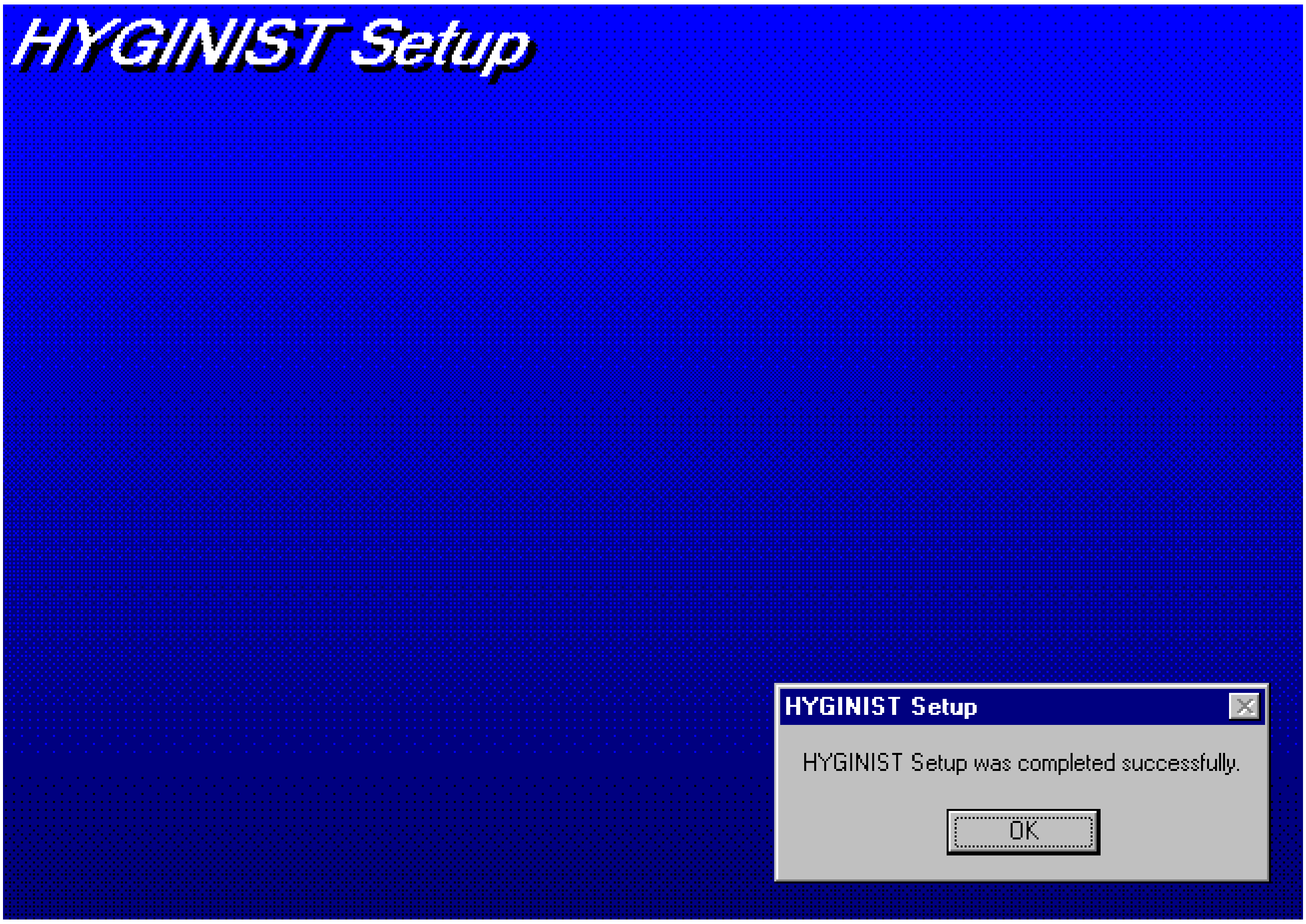
Now click the large button in the upper left corner to install the program in the default directory.
You can change the standard (sub)directory C:\Program files\HYGINIST by clicking the "Change Directory" button on the right site.



All HYGINIST files are copied to the destination directories.



An program Icon to start HYGINIST is set under 'Start', 'Programs'.



The installation is completed successfully! You now can start HYGINIST on your system

Contents - Index

C.5 Start the program

You start HYGINIST for Windows by clicking 'Start', 'Programs'.
Now click on the HYGINIST 

Icon.

If there is no Icon in your menu or on your desktop then do the following:

- By default HYGINIST for Windows is installed in the directory C:\Program files\HYGINIST\
- If you haven't accepted the default directory and don't know were HYGINIST was installed then search HYGINIST for Windows using the Explorer. If HYINIST is found then (double)click on HYGINWIN.EXE.

If this is the first time the program is started after installation took place, then the HYGINIST Login screen appears. If you already filled out Login form then only the Splash screen is displayed for a few seconds before the exposure data management screen appears.

Contents - Index

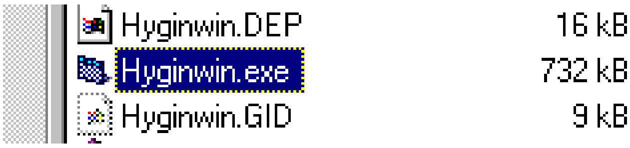
Start, Programs

Contents - Index

HYGINIST Icons

Contents - Index

Explorer



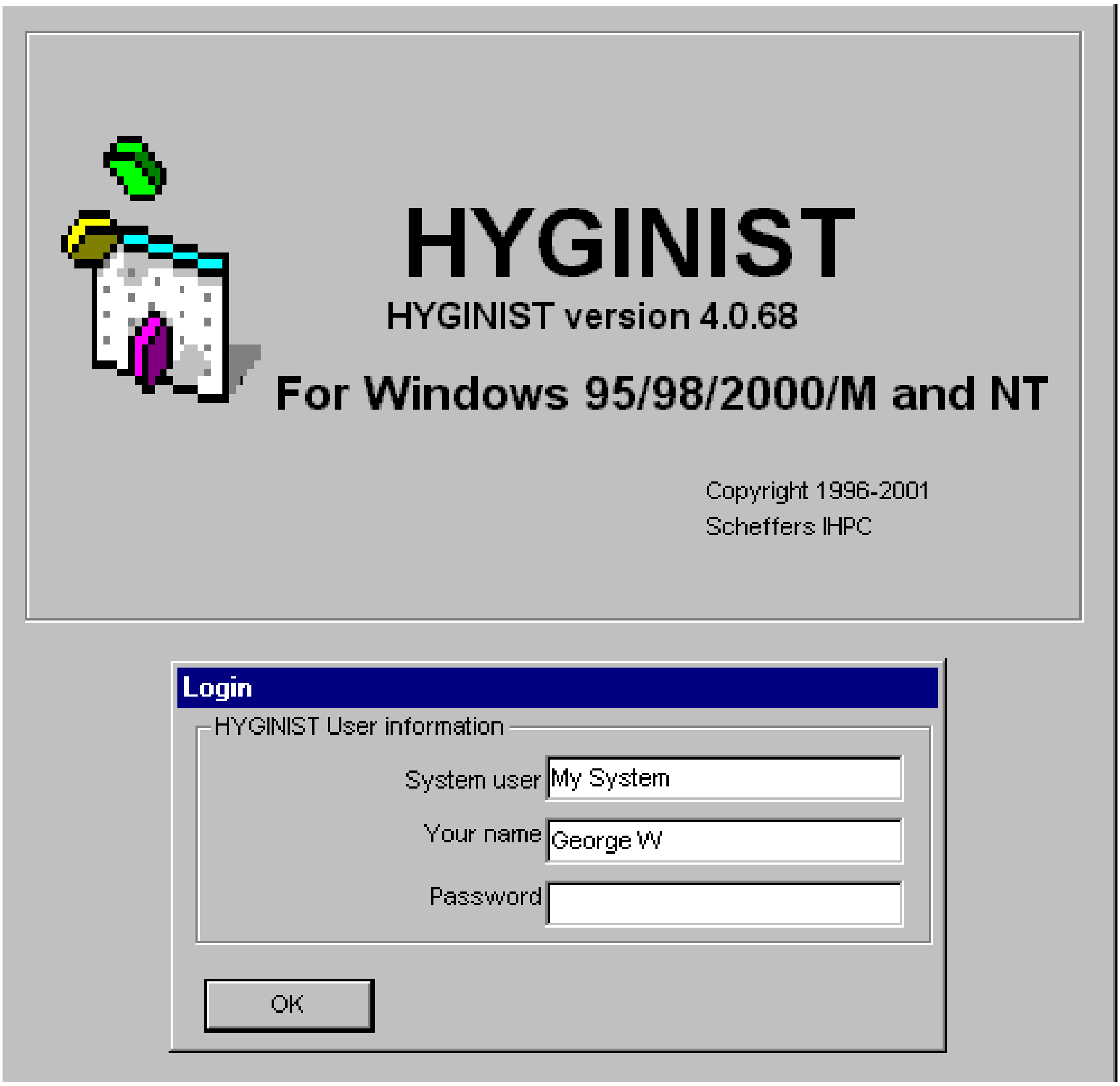
Contents - Index

C.5.1 Enter name and password

A login screen appear. The name of the system user of the computer is displayed

If you have received your password in reply to your registration then fill out:

- your name



- and the password

Press <OK> if you entered all aivallable information.
If the information is correct the program
<Yes> to get a password or <Esc> for direct use of an unregistrered version

How to Register

Choose between a stand-alone or a network licence.
To register fill out the registration form.

When registration form and the "System user" and "Your name" of the Login screen are received and also payment is secured, then a HYGINIST password will be returned to you by email.

Additional information may be found at the HYGINIST homepage

Contents - Index

Stand-alone license

License structure at January 1, 2001. This license structure replaces all preceding
The stand-alone license is meant for use on individual PC's.

The stand-alone version costs EURO 120.- .

For more than five stand-alone packages, special rates are counted.
Please contact Scheffers IHPC for this.

Contents - Index

Network license

Network license structure since January 1, 2002. This license structure replaces all preceding.
The network license is meant for organisations were PC's are linked to a central server and a network director provides the end-users with the software.

The base package (zero licenses) EURO 35,-

Every licence up to ten, per license EURO 85,-
Eleven and more licenses, per license EURO 60,-

So the price for
a single-user network license 85+35= EURO 120,-
10 users 10*85+35= EURO 855,-
20 users 10*60+10*85+35= EURO 1485,-

For more than 50 licenses, special rates are offered.
Please contact Scheffers IHPC for this.

Contents - Index

HYGINIST for Windows registration form

If you decide to register HYGINIST for Windows then please fill out this form and send or fax it to:

Scheffers IHPC
Cramer van Brienenstr 1F
6225 BA Maastricht, The Netherlands
Fax +31(0)842-201756

Your name:

Company:

Shipping address:

City, Zipcode: Country:

Your phone: your fax:

Methods of Payment.
Using your Visa, Master or Euro creditcard;
Please charge to my account (write here the total price):

Card Number:
Name (on card):
Expire date:

From within the Netherlands you can transfer the total amount to the Postbank Account 6729317 of Scheffers IHPC mentioning HYGINIST for Windows.
Vanuit Nederland kunt u het totaal bedrag overmaken op giro 6729317 van Scheffers IHPC, Maastricht onder vermelding van HYGINIST voor Windows.

Product Unit price Quantity TVA Net price
in Euro's (€) € €

Stand-alone or
single user network license € 95,- * .. = ..,-
Clients within Europe
add 19% TVA/VAT/BTW € 18,05 * .. = ..,-

2 to 10 network licenses add € 60,- * .. = ..,-
Clients within Europe
add 19% TVA/VAT/BTW € 11,40 * .. = ..,-

11 and more licenses add € 40,- * .. = ..,-
Clients within Europe

add 19% TVA/VAT/BTW € 7,60 * ..   = ..,-

Total TVA   + €..,- -> ..,-

Total amount   + €..,-

Name: Date

Signature   /   /

The HYGINIST Pasword will be returned if payment is guaranteed and the email with the "System User" and "Name User is received.

Contents - Index

Register HYGINIST

HYGINIST is shareware, so you have the opportunity to fully try it out before you register. Just press the <cancel> button in the Login screen every time you start the program. In the following screen the free trail period will be indicated.

If you want to use HYGINIST without shareware warnings, then you have to do two things:

- fill out the registration form and send or fax it to:

Scheffers IHPC
Cramer van Brienenstr 1F
6225 BA Maastricht
The Netherlands
Fax +31(0)842-201756

- email the "System user" and "Your name" of the Login screen to ihpc@planet.nl

When registration form is received and payment is secured, the HYGINIST password will be returned to you by email.

Contents - Index

C.5.2 How to get a password

If you don't know the password or pressed the <Enter> key before filled it out, then the next message box appears:

- If you want to have a password because you will be using HYGINIST regularly, then press <Yes>. Your email program will be activated in oder to request a pasword.
- The HYGINIST help file will be activated and you will be instructed in how to obtain a password.
- If you have a valid password but you entered before you could fill out the Password textbox then press <No>

- If you just wat to see how the program works then press <Cancel>
- Do you want to know the license agreement then press <Help>

C.5.3 Use as unregistered shareware

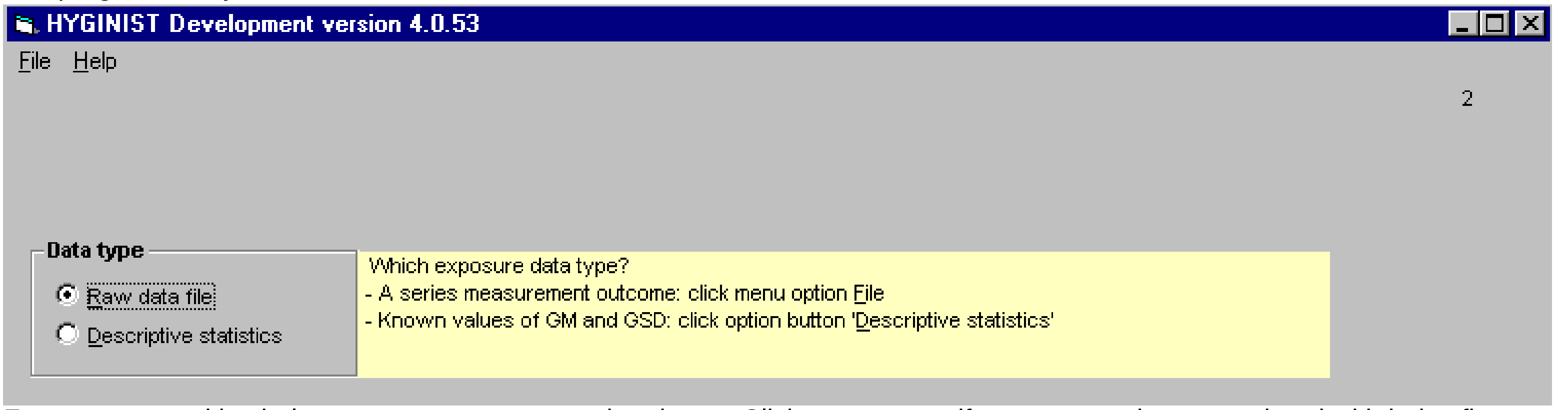
Contents - Index

2. Exposure data management

HYGINIST evaluates two types of exposure data:

- A series of M measurement results , or
- A sample size M and the Lognormal statistical descriptive statistics GM and GSD.

The program always starts with the next window.



For context sensitive help move your cursor over the picture. Click your mouse if your cursor changes to hand with index finger and you get context sensitive help.

this screen offers you the possibility to choose between:

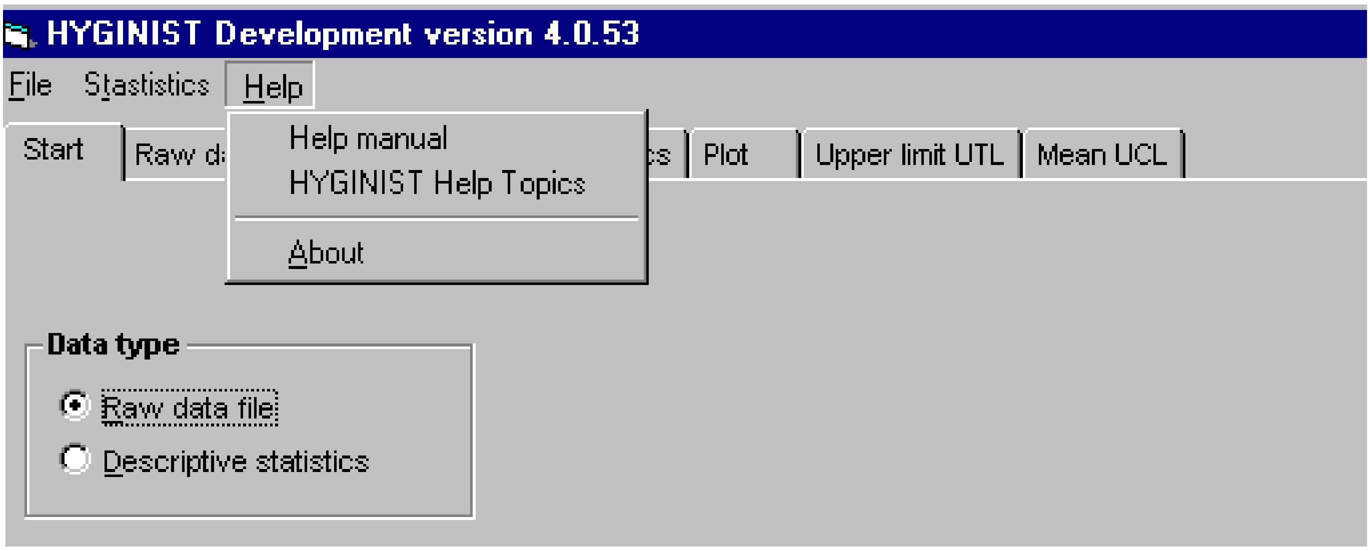
- 'Entering raw data' by keyboard,
- Entering the estimators of the Lognormal 'descriptive statistics' by keyboard,
- Loading a raw data file from disk.

HYGINIST exposure data management further includes:

- Instruction about the requirements that all exposure data should meet (par. 2.1),
- Input of exposure data (par. 2.2),
- ~~Sorting and calculating descriptive statistics (par. 2.3),~~
- ~~Error messages and handling (par. 2.4),~~
- ~~Save, show and edit data files (par. 2.6 through~~ 2.8),
- Exchange files with other database programs (par. 2.9),
- At this moment it is not possible to import exposure data directly form a database or a spread sheet

Contents - Index

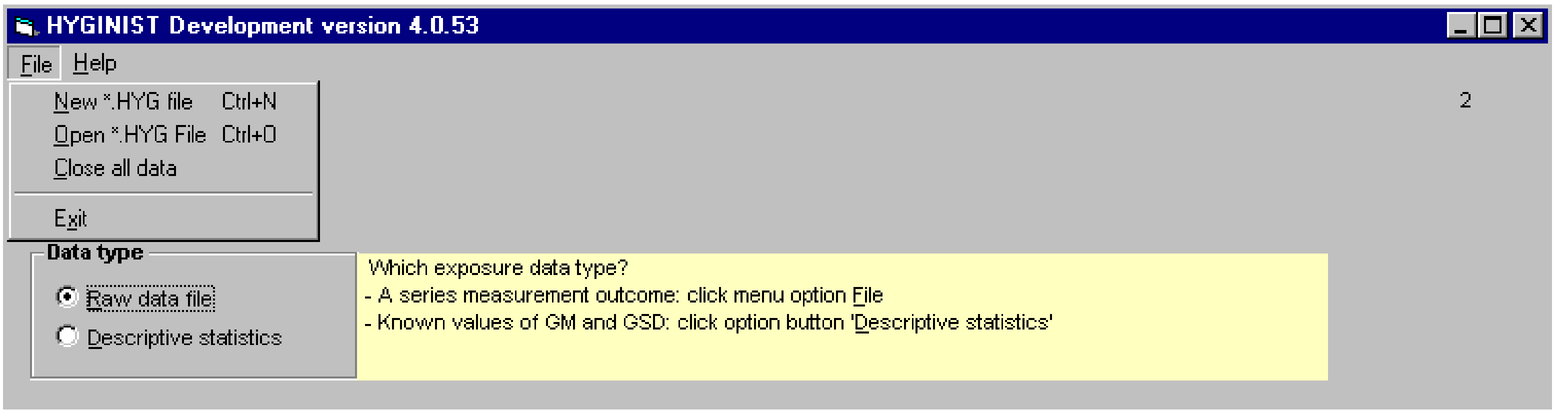
E.2 Online help



Contents - Index

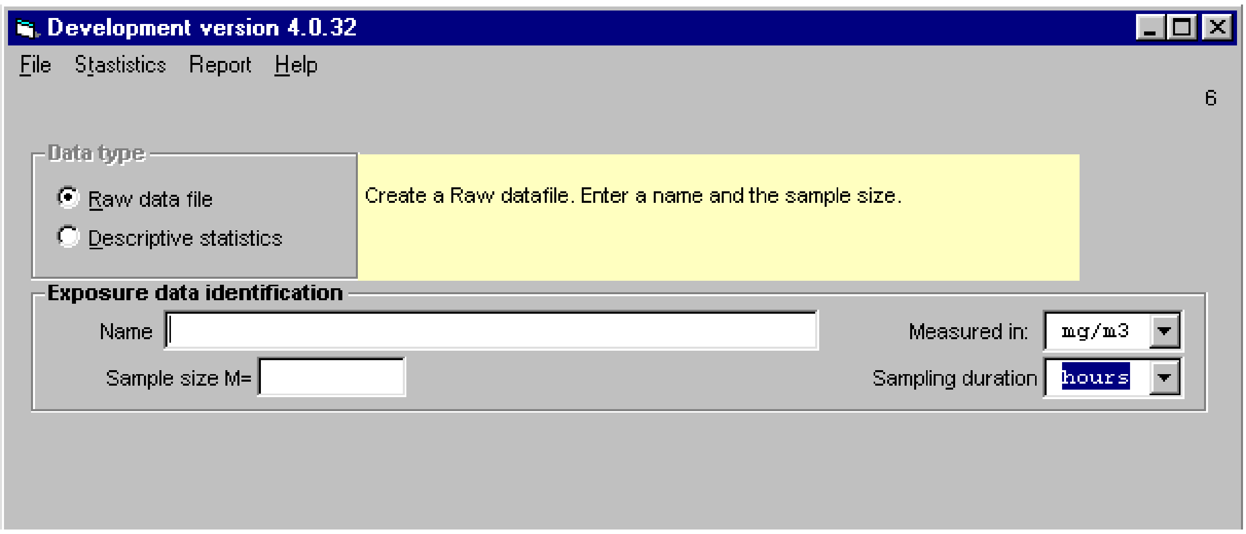
2.2.3 Entering raw data

Click '&File', 'New *.HYG file' to start entering raw data.

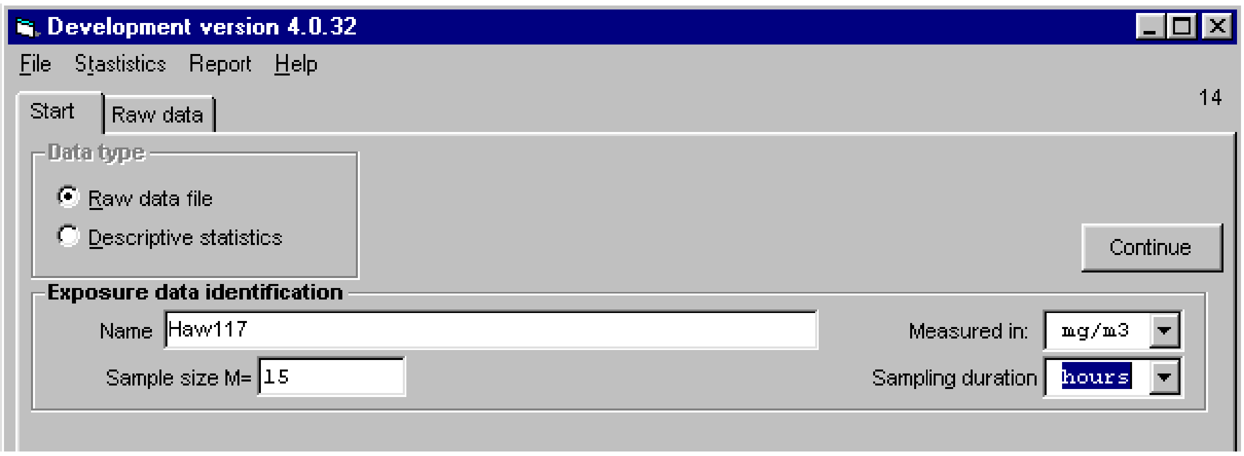


The frame 'Exposure data identification' in which you can fill out :

- a name (obligate)
- A sample size (obligate)
- Units of measurements
- Sampling duration

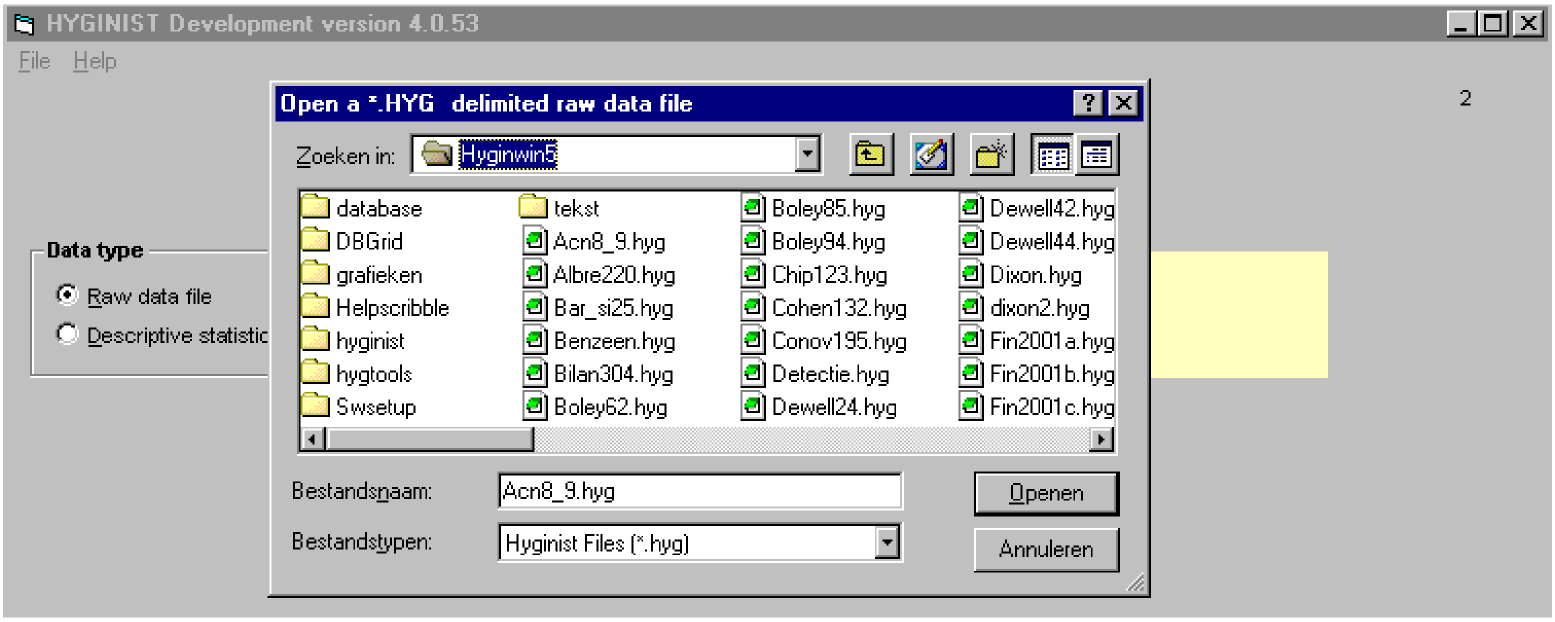


If name and sample size are filled then two tabs appear and a <Continue> button.
Click on tab 'Raw data' or on button 'Continue' to start entering individual measurement outcome (2.2.3.1).



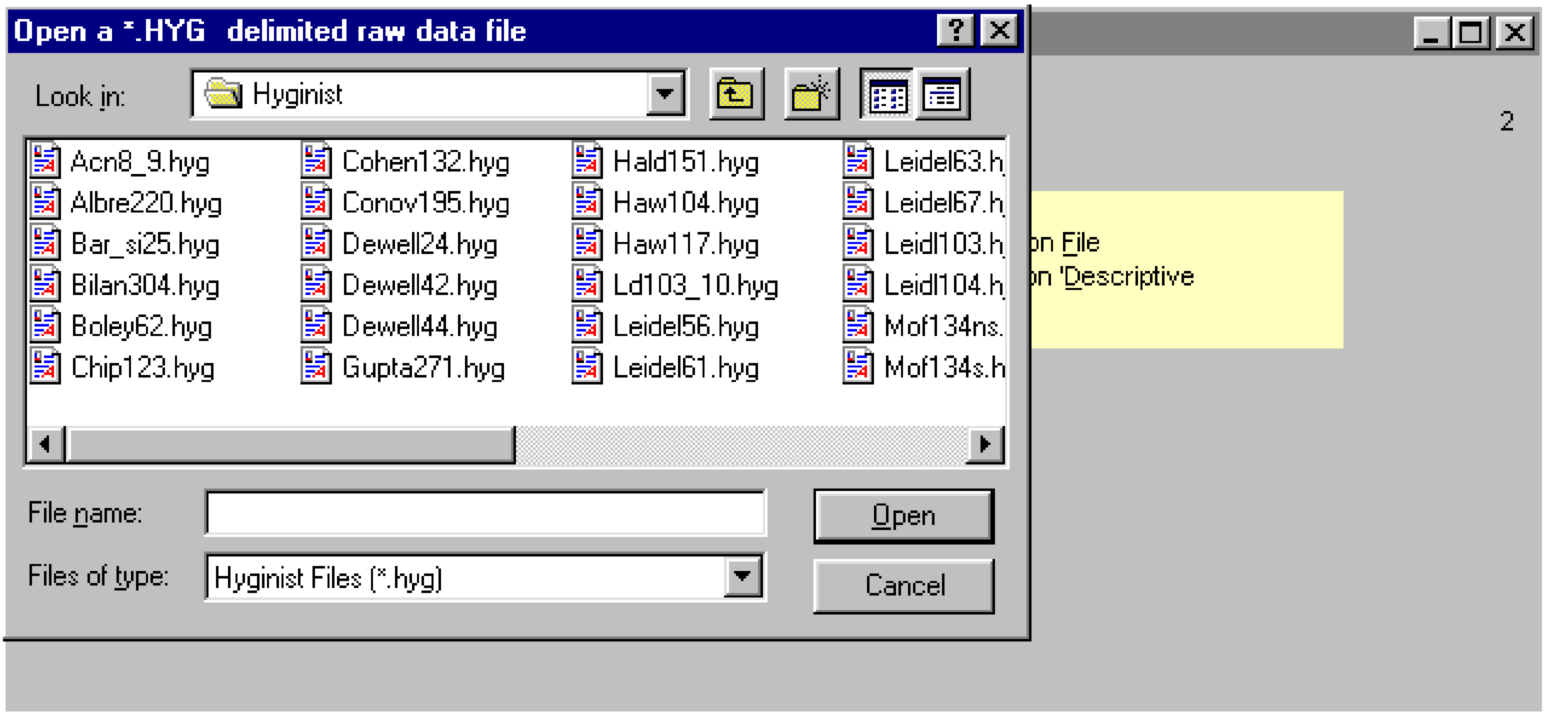
Contents - Index

2.2.4 Loading a file



2.2.4.1 Select a file from disk

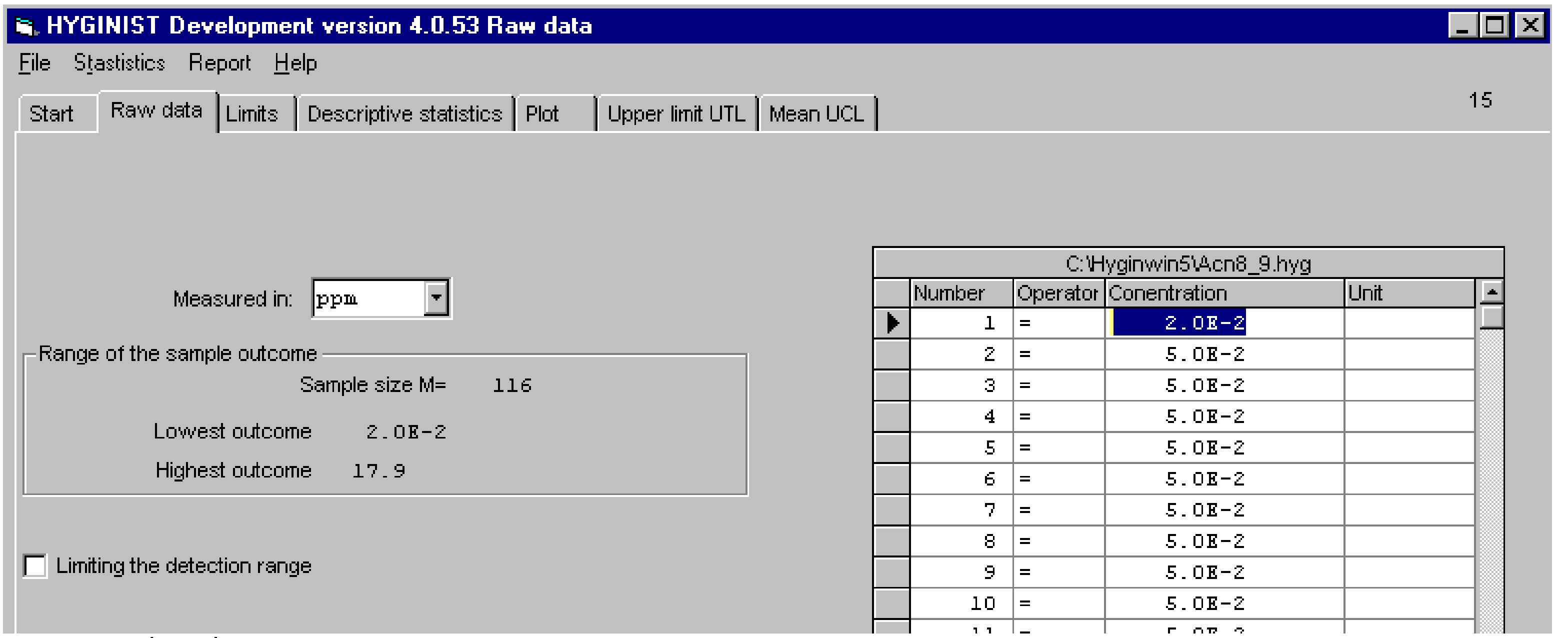
Click '&File', 'Open *.HYG file' to load an existing raw data file.



Contents - Index

2.8.2 Editing raw data

Change the value of a result by clicking in the grid and selecting the value.



Now retype its value
If a series raw data from the current directory contains:
- values outside the range (1E-10 through 1E+10),
- just equals,
- too much variance,

a mesage box apears with an error messages.
This offers you the possibility to correct a file (for example to introduce detection limit values in stead of zeros).

Example 5 Show raw data file

Contents - Index

2.2.1.2 Units of measurement

Enter units of measurement (=dimension):
- to prevent errors in entering a limit value
- as an aid for the report facility,

The dimension characterizes the intensity of the exposure. The sampling method decides on the denominator and the analytical methods on the numerator. Indicator tubes and direct readable devices for gasses and vapors often use 'parts per million' (ppm = ml/m3). For toxicological active gasses and vapors 'weights per volume' (like mg/m3) as the dimension is mostly undesired. The use of these measures originates, however, from the analytical laboratory procedure. According to EN 689 (CEN 1995) the following dimensions for workplace atmosphere should be preferred:

- gasses and vapors ml/m3 (ppm, independent of temperature and pressure) or mg/m3 at 20 °C and 101,3 kPa,
- fibers/m3,
- aerosols (suspended matter) mg/m3

2.2.3.2 Censored data

Treat measurements outside the range of accuracy as follows:

- define an upper and/or lower limit (LL, UL) outside the range of accuracy (ll through ul), so that no outcome lie outside the range of accuracy and inside the limits (LL through ll and ul through UL);
- give results outside the range of accuracy the value of the limit.

Example 3 Range of accuracy

Contents - Index

note 12

There is no single accepted way to handle undetectable levels (Leidel 1977 p 103). The procedure followed here is introduced to get a clear graphical presentation in the probability plot (screen 17).

Contents - Index

Example 3 Range of accuracy

*.HYG Description
HAW104 An 'Acid Operator' in a (hypothetical) Sodiumchlorine plant
is measured for chlorine gas. 10 random $TWA_{8\ hour}$ workplace
air exposure samples were collected (Hawkins 1991 p 104).
The results in ascending order are
<.1, <.1, <.1, .1, .2, .2, .3, .3, .4, .5 PPM.
The lower accuracy limit of the measurement technique is .1 PPM.
Because one result is exactly .1 and three results are <.1,
the lower detection limit is established at .05.
The following data are entered
.05, .05, .05, .1, .2, .2, .3, .3, .4, .5 in a file called HAW104.HYG.

Contents - Index

Example 5 Show raw data file

*.HYG file Description
LEIDL103 To eliminate the fixed background concentration of .1 PPM HF in the file LDL103_10 LEIDL103.HYG, subtract .1 PPM from all data and save the result as LDL103_10.HYG
SCHNE224 Items of which life time under stress is the quality characteristic, are tested (Schneider 1988 p224). The survival time of 15 items from a batch of 50 were 1.6, 2.0, 3.3, 3.3, 3.4, 3.5, 3.8, 4.4, 4.5, 4.5, 4.8, 5.0, 5.8, 5.9, 5.9. (Schneider 1988 Table 7.2.). The remaining thrityfive exceeded the survival time 6.0. Change the default value zero to 6.0

Contents - Index

2.2.2 Assign a name to the exposure data

Every file name should start with a character. The obligate name extension `.HYG` can be omitted. The name length may not exceed 255 characters

If a new file with raw data is created and saved with an existing name, then the data of that file are removed.

If only descriptive statistics will be entered, then using an existing name has no consequences for the raw data in that file.

Contents - Index

2.2.1.1 Sample size

Enter the sample size. If more than two measurements are entered, then examining the shape of the sample distribution is possible.

If the goal of a HYGINIST session is the assessment of:
- The NIOSH non-compliance probability (see par. 5.1.2), or
- The minimum sample size for long-term control (see chapter 7),

and the descriptive statistics GM and GSD will be entered, then any integer between the limits is valid.
More on sample size see 2.2.1

Example 1 Reanalyzing exposure data by adapting the number of degrees of freedom

Contents - Index

3 Plot

The tab <Plot> includes:
- the examination of the Lognormal shape of a sample exposure data, and
- descriptive statistics estimation between limits of accuracy.

Table 4 in paragraph 3.4 helps the user to draw proper conclusions on the shape. A Lognormal shape is essential for an effective extrapolation to general results (Hawkins 1991 p58).

The tab "Plot" presents:
- a plot of the sample on Lognormal probability scale (Screen 17a, see par. 3.1),
- defining the accuracy range of the sample measurements (Screen 17bcd see par. 3.1.2)
- linear regression (Screen 17e, see 3.1.3)
- regression estimators for the descriptive statistics (Screen 17e, see 3.1.4)
- testing the shape (Screen 18, see par. 3.2),
- seven different transformations (Screen 19, see par. 3.3).

If at least three data entries are entered by the keyboard or loaded from a file, then this tab is visible.

The exposure data in example 9, example 10 and example 11, and the corresponding figures, are used throughout the manual to illustrate the different aspects on Lognormal extrapolation and inference in the following chapters.

Contents - Index

3.4 Decision scheme

The following procedure, based on both industrial hygiene and statistical considerations, helps to decide if the Lognormal model is indeed the most effective model to describe this series exposure data (A.4.1 and proposition 1).

1 establish decision criteria, based on strategic 3.4.1 or compliance 3.4.2 considerations,

2 examine the assessment strategy, in particular on:

- homogeneity in the population(s) under study (see A.2),
- autocorrelation, if $GSD_{TWA\ 8\ hour}<2$ (Spear 1986, Francis 1989, Rappaport 1991),

3 examine the sample series on outliers and remove them from the sample if there are grounds for doing so (see B.3),

4 establish the accuracy range and set upper and/or lower limits in Screen 17bcd (see 3.1.2), if raw data exceed the accuracy range (see B.4.3),

5 examine the conformity with the Lognormal distribution using Table 4,

6 test the results from 5 with the decision criteria in 1 and decide on the conformity. If you decide that the conformity is not rejected, then continue with extrapolation and compliance testing.

Table 4 Lognormal goodness-of-fit decision scheme

Decide on situations not covered by Table 4, through an evaluation of the exposure assessment strategy.

3.4.1 Strategic considerations

Lognormal goodness-of-fit:

- is of minor concern in range finding research,
- should not be rejected in long-term compliance control and/or routine monitoring programs,
- is important in first establishing similar exposure groups (see A.2.3).
- is important if the NIOSH method for long-term TWA control is used (see 5.1.2).

If the observed shape is not in compliance with the goal of the exposure assessment, then both sampling plan and the strategy should be reconsidered.

Contents - Index

3.4.2 Compliance considerations

Goodness-of-fit is:

- important if the limit H lies in the central region of the exposure distribution, that is:
- $C_1 \leq H$ and $C_M \geq 0.1*H$, or
- $GM<H$ and $GM*GSD^2>H$.
- less important if robust extrapolation distributions are used, like the (Noncentral) Student
- of only academic importance if $GSD<1.4$, or
- if limit H is an outlier in the exposure distribution, that is:
- extremely high, $GM \geq H$ or $C_{M/2} \geq H$, or
- extremely low, $GM*GSD_2 \leq H$ or $C_M \leq 0.1*H$,

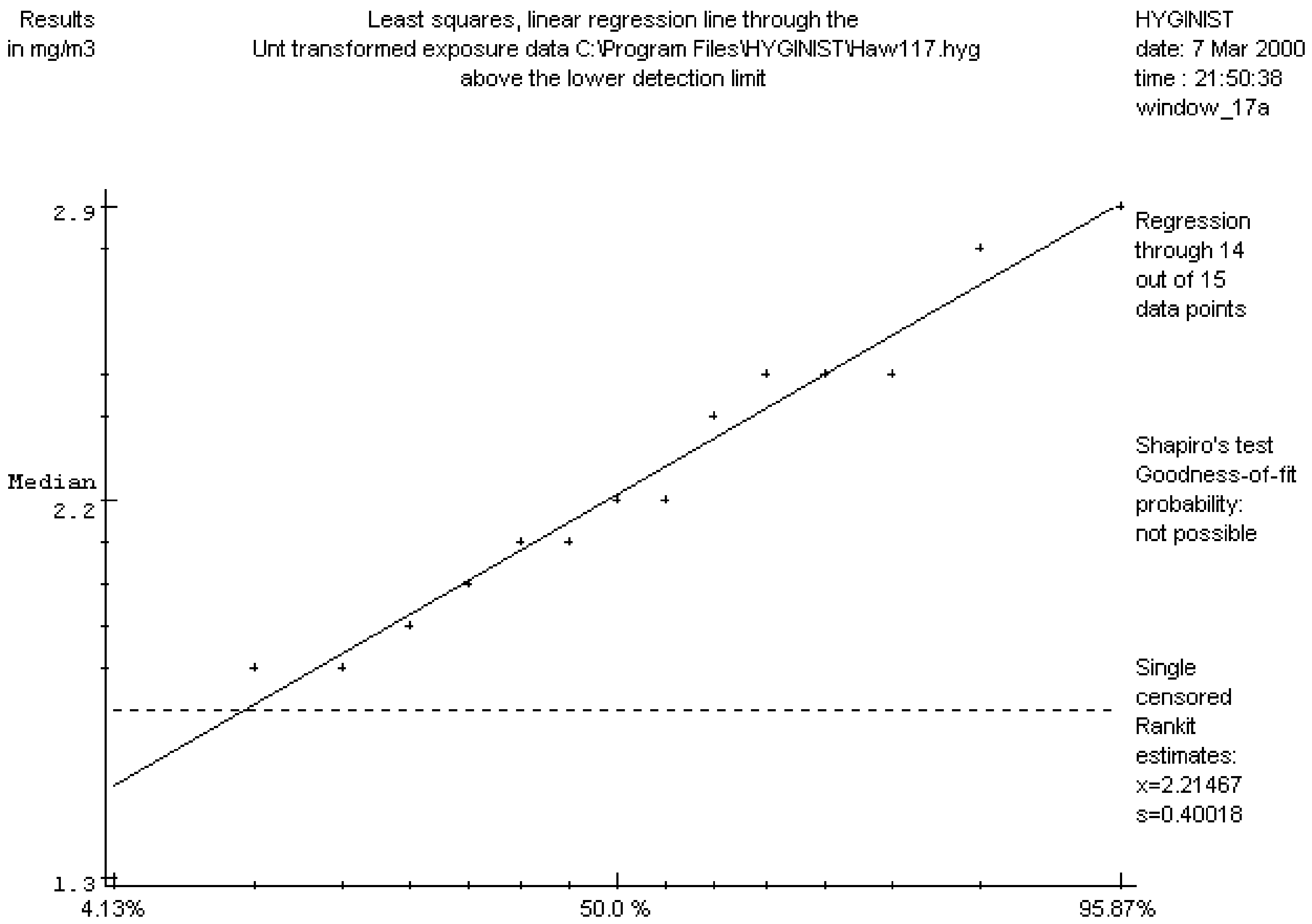
Figure 11 15 Untransformed TWA8 hours Figure 12 Total dust 12 TWA8 hours dust (HAW117) (BOLEY62)

Example 16 Conformity with the Lognormal model. 28 files with A(W)logD5%

Example 17 Conformity with the Lognormal model based on the graph. Complete files with A(W)log<5%

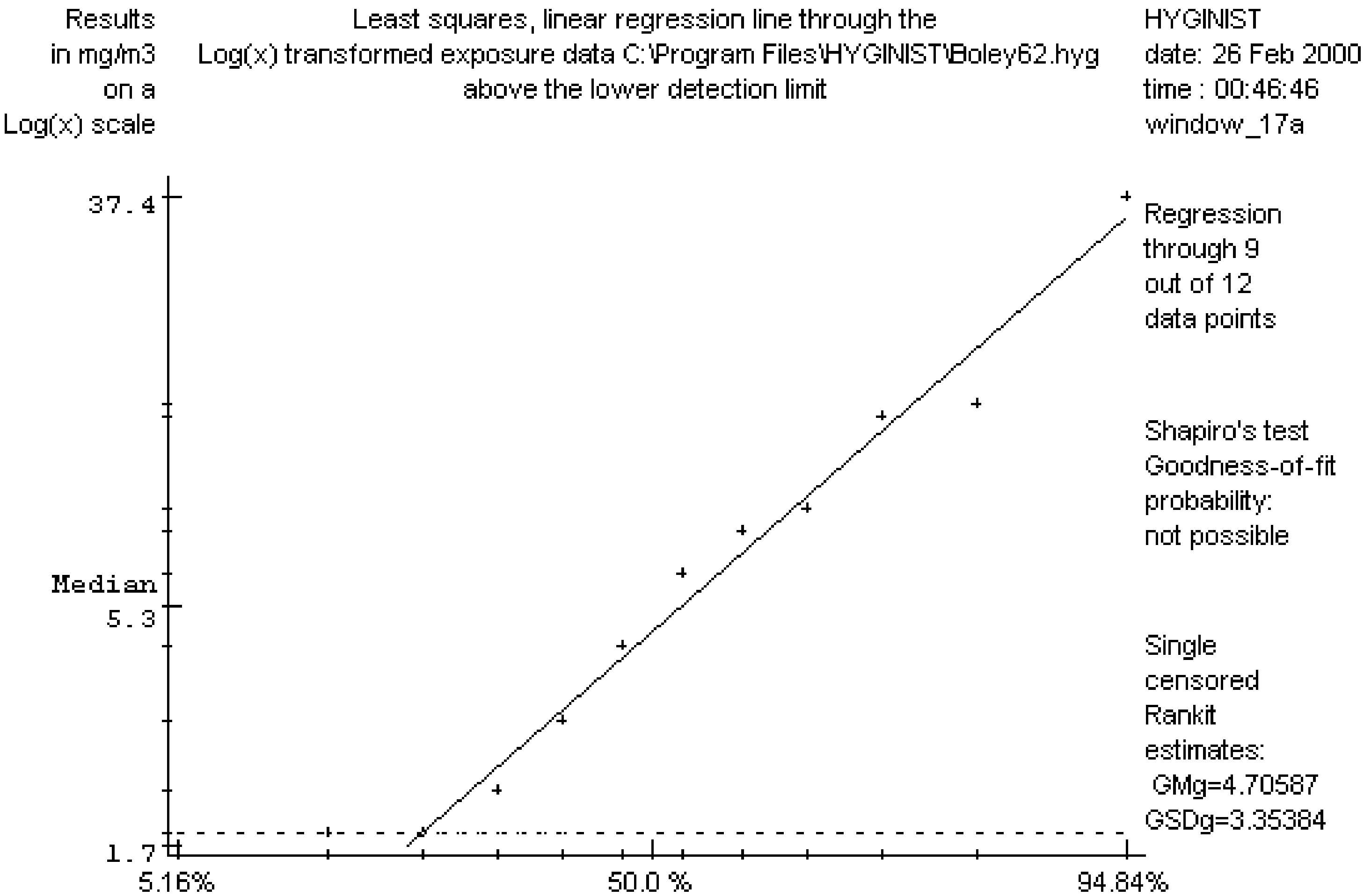
Contents - Index

Figure 11 15 Untransformed TWA8 hours dust (HAW117)



Contents - Index

Figure 12 Total dust 12 TWA8 hours (BOLEY62)



Contents - Index

Example 16 Conformity with the Lognormal model. 28 files with A(W)>=5%

*.HYG file Description of the exposure data Shapi A(W) Lognormal
name ro W in % Conformity
LEIDEL67 10 one worker $TWA_{8\ hour}$ PAS Dioxane (Figure 1) 0.99 99.5 Good
LEIDL104 24 $TWA_{8\ hour}$ PAS Methyl Methacrylate (Figure 21) 0.99 97.3 Perfect
LEIDEL63 5 one worker, one shift $TWA_{10\ min}$ PAS CS_2 0.99 94.8 Good
LEIDEL61 35 random, momentary workplace Ozone read off's 0.98 92.1 Good
POST10P5 4 one worker, one press $TWA_{8\ hour}$ PAS Styrene 0.98 90.4 Good
RANDOM20 2000 Standard Lognormal deviates 0.99 87.5 Good
BAR_SI25 6 Carbon monoxide grab samples 0.94 64.4 Rejected
DEWELL44 11 SEG $TWA_{8\ hour}$ PAS welding [MIG] fume 0.94 51.4 Acceptable
POST08P4 5 one worker, one press $TWA_{8\ hour}$ PAS Styrene vapour 0.92 50.8 Acceptable
LD103_10 12 corrected grab sample HF correction (Figure 13) 0.94 50.7 good*
COHEN132 20 maximum flood levels (Figure 14) 0.96 49.4 good*
BILAN304 5 $TWA_{8\ hour}$ Methylene bisphenyl isocyanate. (Figure 9) 0.91 43.7 Acceptable
LEIDEL56 8 $TWA_{20\ min}$ PAS Ethyl alcohol vapour (Figure 10) 0.92 41.8 Acceptable
CHIP123 10 $TWA_{8\ hour}$ PAS lead dust 0.93 40.0 Acceptable
MOF134S 16 $TWA_{10-30\ min}$ Methoxyfluorane. Scavenged (Figure 20) 0.90 37.2 Acceptable*
HAW117 15 SEG, random $TWA_{8\ hour}$ PAS total dust 0.94 35.7 Acceptable*
DEWELL24 8 one worker $TWA_{8\ hour}$ PAS on foundry respirable dust 0.91 34.3 Rejected
SOLV198 45 housepainter $TWA_{8\ hour}$ PAS Hydrocarbon (Figure 2) 0.97 33.9 Acceptable
BOLEY62 12 one worker $TWA_{8\ hour}$ PAS total dust (Figure 12 page 38) 0.93 33.4 Acceptable*
ALBRE220 4 PAS $TWA_{8\ hour}$ Aspartame total dust 0.87 29.2 Acceptable
POSTB11I 14 $TWA_{8\ hour}$ PAS Styrene vapour (Figure 3) 0.93 29.0 Rejected
X07-10 35 concentrations of chemical X 0.94 17.6 Rejected
GUPTA271 10 mice, survival (Figure 26) 0.89 14.8 Acceptable*
HAW104 10 random $TWA_{8\ hour}$ PAS Chlorine (Figure 8) 0.87 10.1 Rejected
POST1008 combination of POST08P4 and POST10P5 0.85 7.6 Rejected
MOF134NS 16 $TWA_{<30\ min}$ Methoxyfluorane. Nonscavenged (Figure 19) 0.90 7.6 Acceptable*
DEWELL42 10 $TWA_{10\ min}$ PAS on formaldehyde 0.86 6.6 Rejected
LEIDL103 12 stationary, grab sample HF, one hour sequence 0.87 6.6 Rejected
*conformity decision after censoring

Contents - Index

Example 17 Conformity with the Lognormal model based on the graph. Complete files with A(W)log<5%

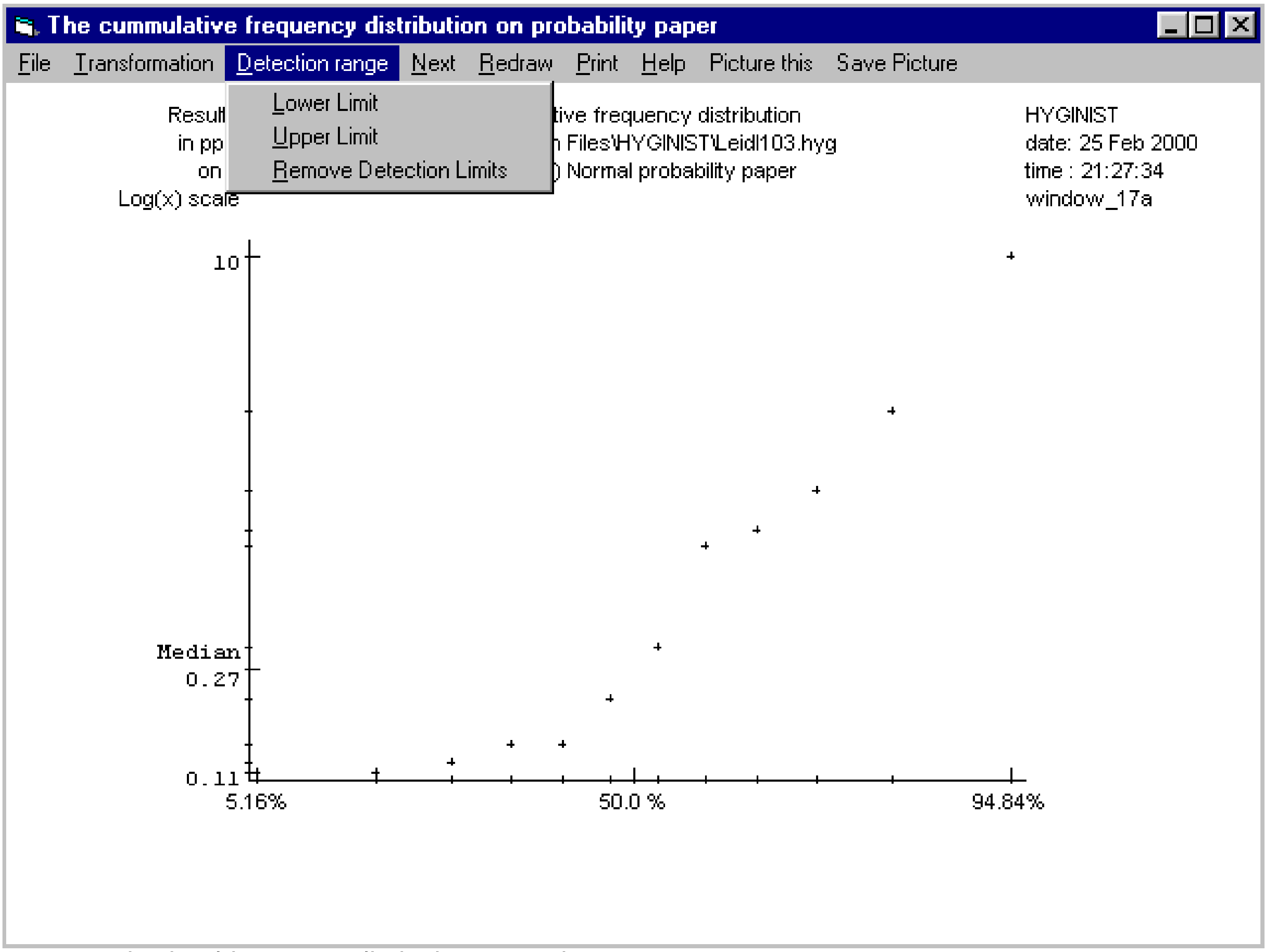
*.HYG file Description of the exposure data Shapi A(W) Lognormal
name ro's W in % Conformity
SCHNE70 failure distance of 96 locomotives (Figure 6) 0.62 1.27D-37 very good
SCHNE224 survival of 50 items under stress (Figure 5) 0.54 8.6D-16 Good
OWEN716 15 grab sample airborne Chlorine (Figure 4) 0.83 0.97 Acceptable

Contents - Index

3.1.2 Accuracy limits

With menu option "detection limits", upper and lower limits can be introduced and removed. A accuracy range prevents that outlier data that cannot be excluded from the sample, bias the estimation of the location and the variance. Using the data within the range of accuracy and the corresponding rankit values:

- the shape is examined (3.1.3),
- the descriptive statistics are estimated (3.1.4).
- See par. 2.2.3.2 on how to enter censored exposure data.



Screen 17b Should accuracy limits be entered?

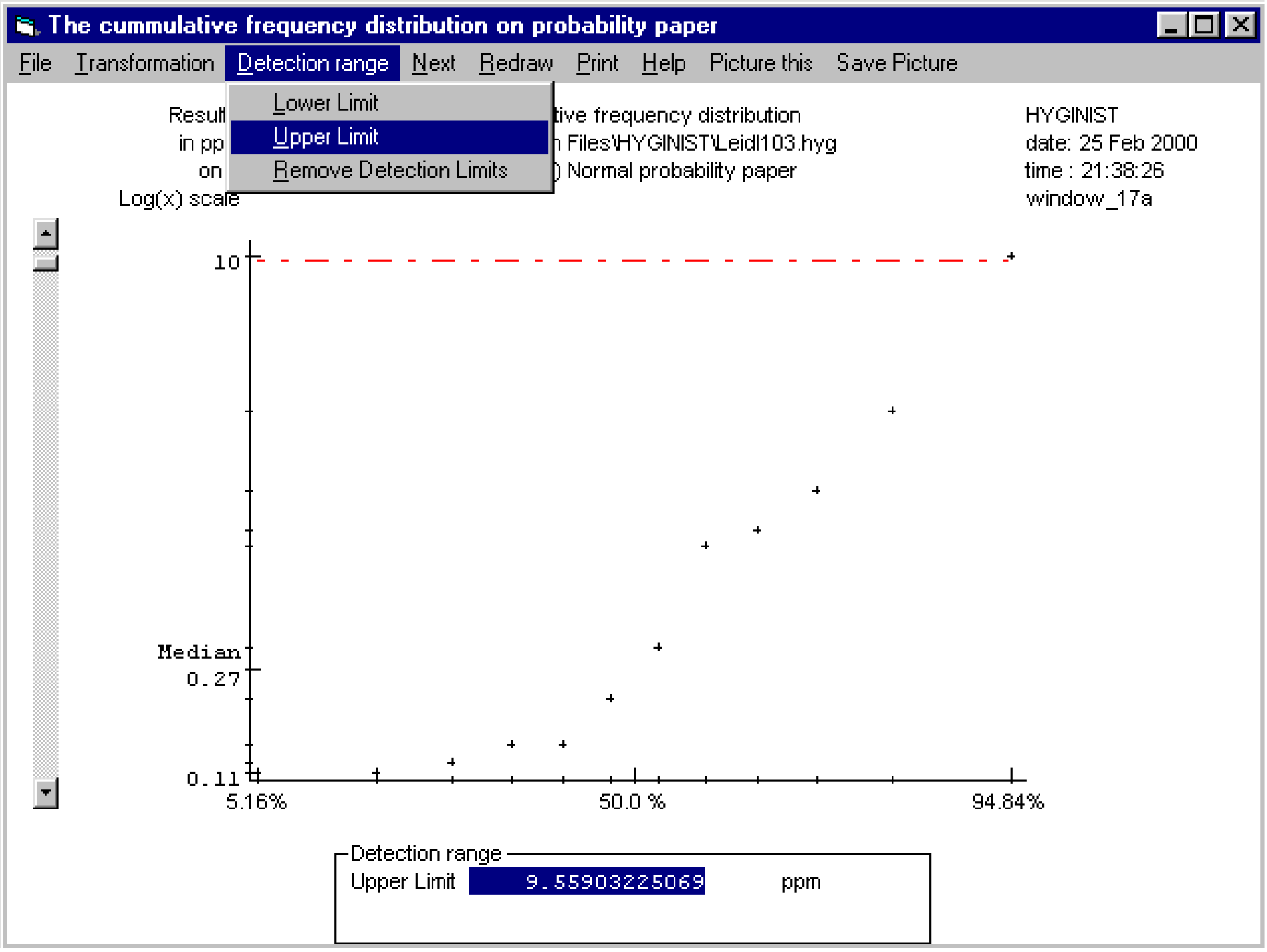
The range of accuracy is always limited:

- All air sampling- and/or analytical methods do have a detection upper and lower limit. TWA8 hour PAS total dust measurements have an accuracy range of about 1.8 to 15 mg/m3 as will be shown in the examples. For aerosols and fibers PAS sampling accuracy limits are narrower than for and gases.

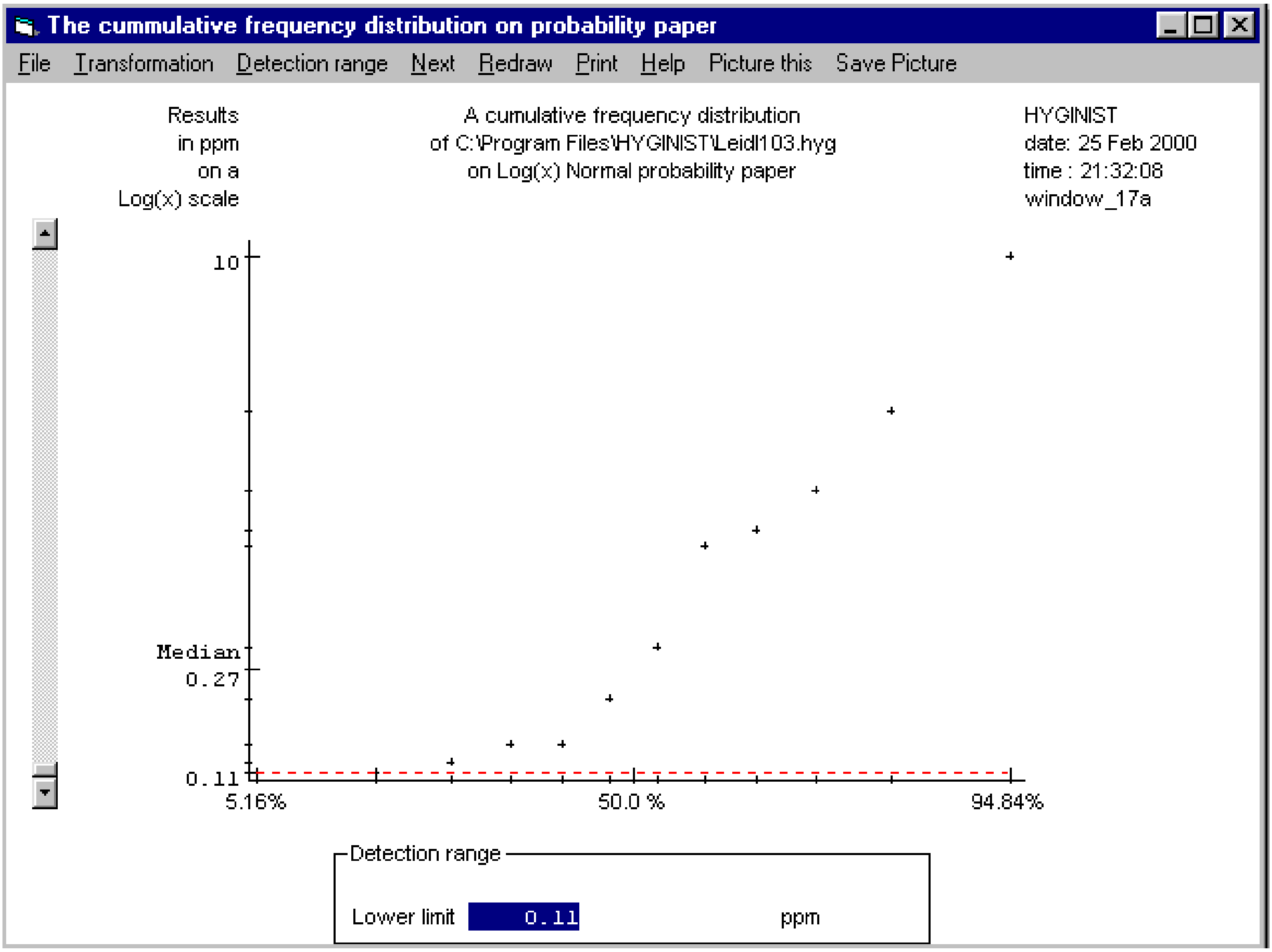
- Non-traceable errors in the strategy execution, the sampling, analysis and/or report procedure, or just by chance, can cause one or more of the extreme result(s) to behave like outliers. To prevent bias on the estimators without eliminating these data, upper and lower limits should be introduced.

Almost all Health and Safety organizations provide analytical methods for the determination of workplace air contaminants (AIHA, ISO, CEN, DFG, EPA, HSE, WHO etc.). They are often based on the manuals from OHSA (1985 and 1987) and NIOSH (Eller 1984). Voice (Dräger), Supelair (Supelco) and DOHS-Base (DOHSBase VOF) are data retrieval, Windows computer programs with industrial hygiene limit values and the detection range of the corresponding measurement methods.

A value of 1.1*C(M) is displayed as default upper limit in Window 17c.



A value of 0.9*C(M) is displayed as default upper limit in Window 17d.



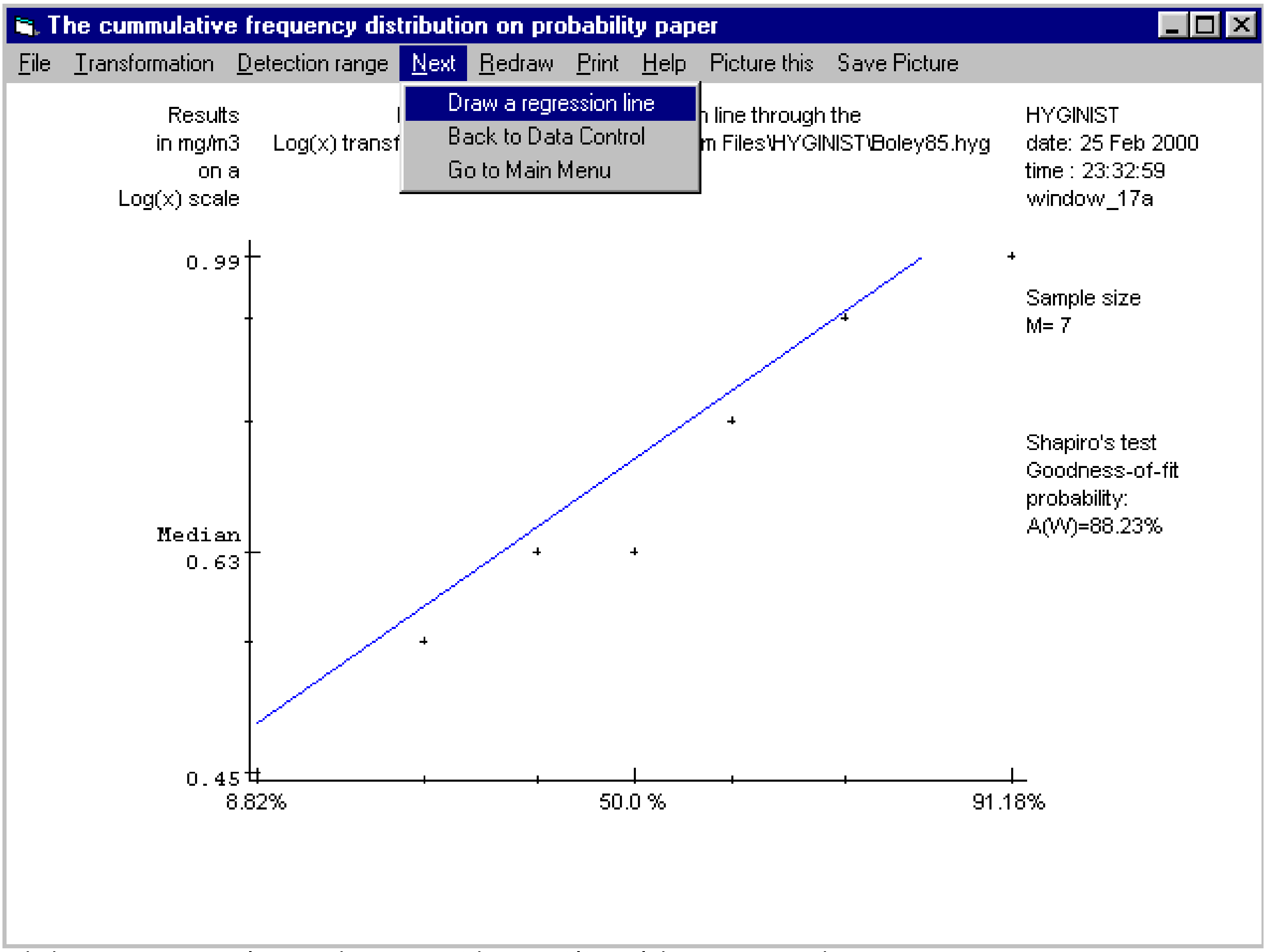
Example 10 [Upper limit]

Example 11 [Lower limit]

Contents - Index

3.1.3 Linear regression

Screen 17e, displaying the linear regression line, is specifically made for:
- Graphically examining the goodness-of-fit,
- Selecting the most effective transformation,
- Making reports



Window 17e Lognormal regression. 7 TWA8 hour ZnCl2 total dust on one worker

The linear regression line is calculated by using the method of least squares (Snedecor 1980 chapter 9). If:
- at least one limit is entered, or
- the transformation is "untransformed",

then Window 17e shows the number of data pairs in the regression ("Regression through #### out of #### datapoints") in the left margin. For the Lognormal case the upper limit $C_{U\%}$ can be estimated by:

$$C_{U\%} = GM_g \times (GSD_g)^{Z_{U\%}}$$ Formula 3-1

With GMg and GSDg calculated using the Formula 4-10 and Formula 4-9.

Contents - Index

3.1.4 Regression estimators for the descriptive statistics

If:

- at least one limit is entered, and
- the transformation is logarithmically or untransformed,

then Screen 17f shows in the right margin:

- the number of data pairs in the regression ("Regression through #### out of #### datapoints"),
- the censoring type ("Single" of "Double") to indicate that one or two limits are introduced,
- the rankit estimators for location and variance based on the data pairs in the regression if the data are logarithmic transformed (GM_g and GSD_g) or untransformed (x_{cg} and s_{cg})

The use of the rankit regression estimators is recommended:

- if a fixed background concentration exceeds the lower variance (see Leidel 1977 p103),
- if the strategy used is based on it (Leidel 1977 p107, WHO 1984, Travis 1990),
- in reanalyzing historical references where it was used.

See B.3.1. for the algorithms of GM_g and GSD_g and their properties.

Example 12 Rankit versus direct estimators

Example 12 Rankit versus direct estimators

To show the effectivity of the rankit estimators a lower limit of 30 PPM was introduced in the 10 $TWA_{8\ hour}$ (LEIDEL67.HYG) PAS Dioxane sample (Figure 1). Because this limit is between the default .9*C_1 and C_1 the rankit estimators are calculated from the complete sample: GM_g=78.4 and GSD_g=1.67. GSD_g is about 2% higher than the direct estimator GSD=1.63 (see Example 17).

Example 10 Upper limit

*.HYG files Description

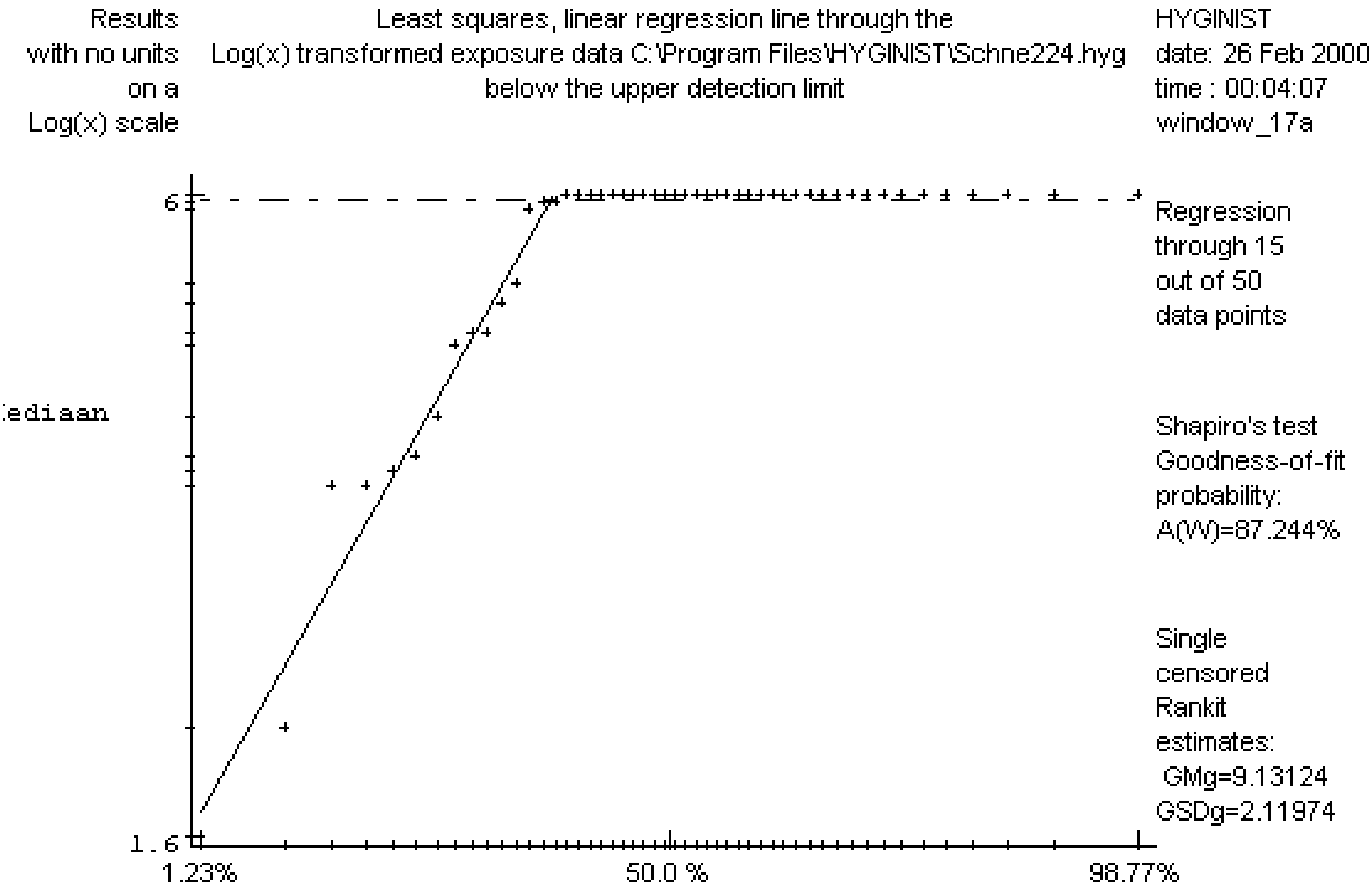
SCHNE224 The survival of items under stress is Lognormal distributed (Schneider 1986 page
224). When the (accelerated) lifetime is above 1 time units (TU) an item conforms
the quality limit. 35 items out of a sample of 50 survived 6 time units (TU).
The results with an upper limit at 6 TU are plotted in Figure 5.

SCHNE70 Schneider (1986 example 4.3 page 69) used a sample of 96 locomotives with 37
failures before the (type I) upper censor point of 135,000 miles, assuming a
Lognormal distribution. This seems to be justified by the rankit Figure 6.

OWEN716 In order to classify the TWA8 hour exposure using 15 grab sample, detection tube
Chlorine concentrations both upper (9 PPM) and lower limit (.25 PPM) should be
used (Owen 1980 p716). See Example 19 for more files with upper limits.

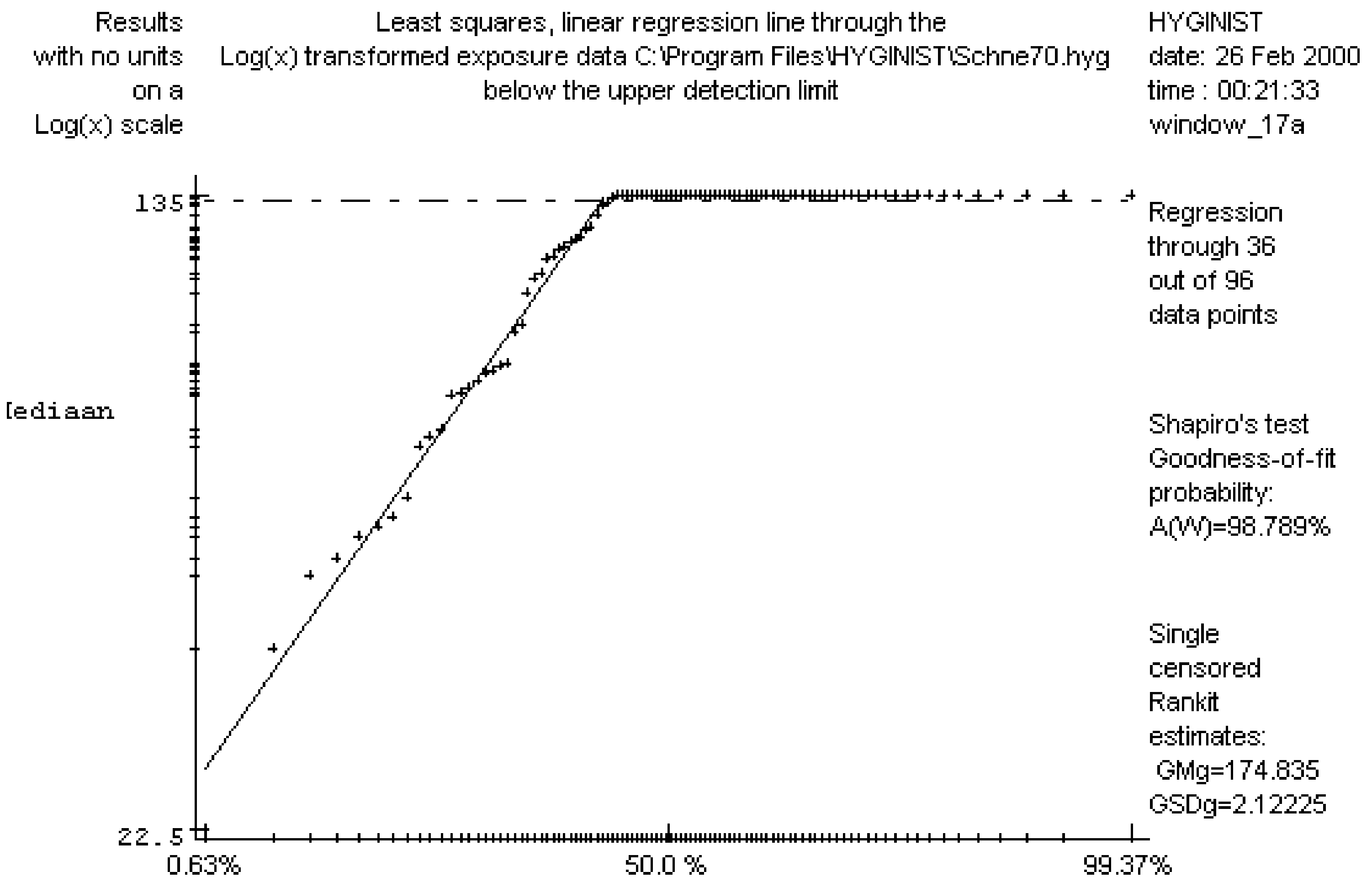
Contents - Index

Figure 5 50 Items under stress (SCHNE224)



Contents - Index

Figure 6 The mileage of 96 locomotives (SCHNE70)



Contents - Index

Example 19 df, GMg and GSDg of censored Lognormal samples

*.HYG file Description of the exposure data Sample units LL UL uncen-df GM_g GSD_g
Name Size sored
M M`
BOLEY62 $TWA_{8\ hour}$ PAS papermill total dust 12 mg/m^3 2.0 - 9 9 4.71 3.35
COHEN132 maximum flood levels 20 10^6 0.27 - 18 18 0.40 1.34
ft^3/sec
DEWELL24 $TWA_{8\ hour}$ respirable dust in 8 mg/m^3 1.3 2 6 6 1.22 1.56
foundry
DEWELL44 $TWA_{8\ hour}$ PAS welding [MIG] fume 11 mg/m^3 - 18.0 9 9 11.12 1.74
GUPTA271 mice survival after inoculation 10 days - 65.0 7 7 56.0 1.24
with tuberculosis
HAW117 $TWA_{8\ hour}$ PAS total dust 15 mg/m^3 1.7 - 14 13 2.19 1.19
LEIDL103 stationary, grab sample, 12 PPM 0.14 - 7 8 0.29 8.27
one hour sequence HF
OWEN716 grab samples airborne Chlorine 15 PPM 0.25 9.0 8 10 0.88 6.81
SCHNE224 survival of items under stress 50 TU - 6.0 15 30 9.1 2.12
SCHNE70 failure distance of locomotives 96 10^3 - 135.0 37 65 174.2 2.12
miles

SCHNE224.HYG the number of degrees of freedom is established as df=2*15=30 [because (50+15)/2-1D2*M]. All others df=(M+M`)\2-1.

Example 11 Lower limit

*.HYG file Description

LEIDL103 Hydrogen Fluoride exposure near a control panel in a production unit was measured
using a stationary grab sampler with a one-hour sequence (Leidel 1977 page 103).
The results are 0.11, 0.11, 0.12, 0.14, 0.14, 0.21, 0.33, 0.8, 0.91,
1.3, 2.6 and 10.0 PPM. The cumulative frequency distribution lacks
Lognormality in the left tail (see Screen 17a). Continued in Example 20.

HALD151 The diameter of population rivet heads follows a Gaussian distribution with
μ=13.426 mm and σ=0.111 mm (Hald 1952 p148). A device with a lower detection
limit of 13.4 mm was used to measure a sample of 500 rivet heads (.HYG).
190 diameters were below the detection limit.
The shape of the untransformed results are displayed in
Figure 24 (grouping interval 0.05 mm). Because of the small scale,
the shape of the sample conforms to both the Normal and Lognormal distribution.

Contents - Index

3.1 Graphical examination

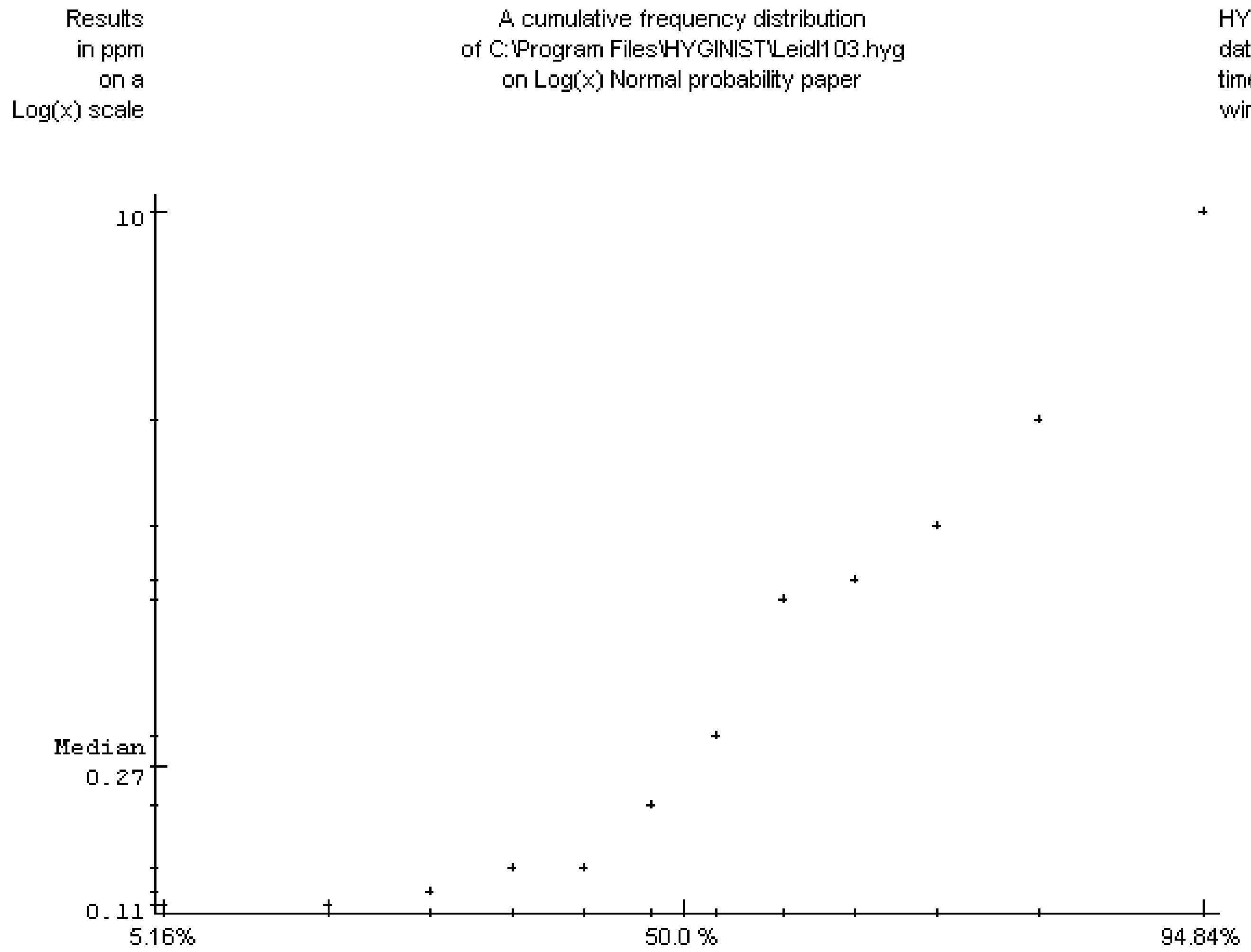
Graphical examination includes:

- a Rankit plot (Screen 17a, par. 3.1.1),
- the assessment of the accuracy range (optional, Screen 17bcd, see 3.1.2),
- a linear regression fit (Screen 17e, see 3.1.3) within the accuracy range,
- the estimation of the descriptive statistics (Screen 17ef, see par. 3.1.4 and 3.1.5),
- plots can be captured and saved, see E.3

Contents - Index

3.1.1 The rankit plot

In the picture box the ascending ordered concentrations are plotted against their expected places (see B.1) within a normal distribution. The (un)transformed concentrations are displayed on the vertical Y-axis. For more information on normal probability plots see B.1.1.



Screen 17a Probability plot of exposure data on rankits scale

The median is displayed in Screen 17a if its value falls within the range of 5-95% of the vertical axis.

Rankit plots are made from all exposure data described in example 9, example 10 and example 11. Example 19 contains censored data files. How the rankit plot is used in the evaluation of shape is explained in paragraph 3.4.

Example 9 Description of rankit plots 1 through 4

Figure 1 Ten TWA8 hours Dioxane (LEIDEL.67) Figure 2 45 TWA8 hours total hydrocarbons among maintenance painters (Solv198)

Figure 3 14 TWA8 hour Styrene on inlayers working on different presses (POSTB11I) Figure 4 15 grab sample airborne chlorine concentrations (OWEN716)

Contents - Index

Example 9 Rankit plots

*.HYG file

LEIDEL67 This example (Leidel 1977 p67) is used in this manual to explain the NIOSH method for the calculation of geometric mean of long-term exposure and the use of the probability of non-compliance when deciding whether to install engineering control (see 5.1.2). On 10 different days in a period of 6 months an employee's $TWA_{8\ hour}$ Dioxane (H=100 PPM) concentration was measured (see Figure 1)

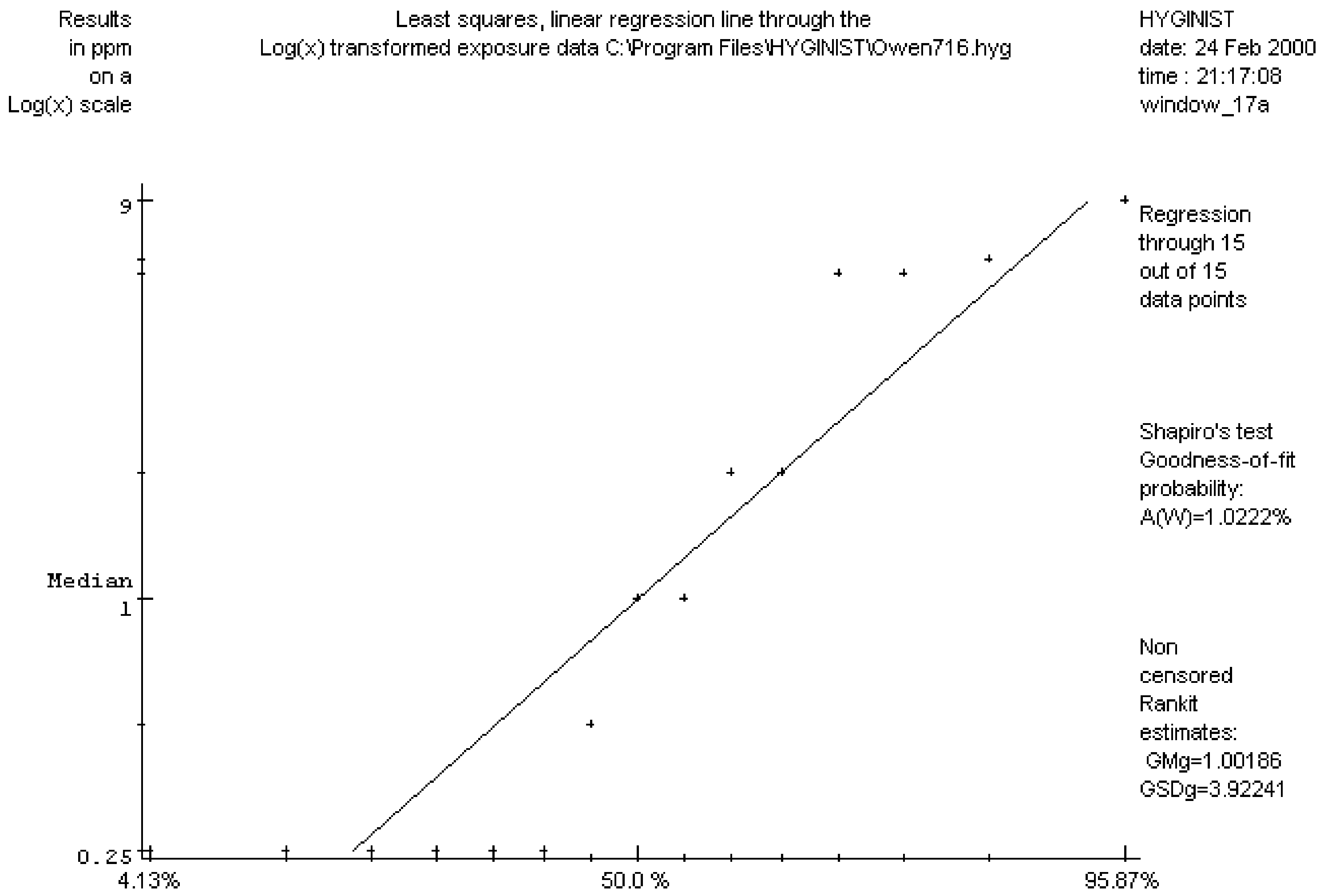
SOLV198 In a Health Hazard Survey among maintenance painters a stratified sampling programme was performed in order to estimate
the current range of total Hydrocarbon exposure (in mg/m^3)
during solvent based paint rolling and spraying (Scheffers 1987).
45 $TWA_{8\ hour}$s were measured.
Although there was no randomisation by purpose, the goodness-of-fit with the Lognormal model is appropriate (see Figure 2).
The two-sided tolerance limits ($1-2\alpha$) are compared with the one-sided tolerance lower ($1-\alpha$) and upper (α) limit (continued in Example 31).

POSTB11I Among a group inlayers in a fibreglass reinforced polyester product manufacturing plant (Post 1989, 1991)
14 $TWA_{8\ hour}$'s styrene vapour were measured.
Figure 3 shows the characteristics of a mixed population situation:
two curved lines linked at an angle.
The study report confirms that the five lowest measurements were taken at a different press than the highest nine.
To compare both populations see Example 39.

OWEN716 At an industrial site in the US, 15 grab sample airborne Chlorine concentrations were gathered over a working day (Owen 1980 p716)
in order to classify for the federal $TWA_{8\ hour}$ standard of 1 PPM.
The results are 0.25, 0.25, 0.25, 0.25, 0.25, 0.25, 0.50, 1.00, 1.00, 2.00,2.00, 6.00, 6.00, 6.50, 9.00 PPM.
The goodness-of-fit is poor (see Figure 4).

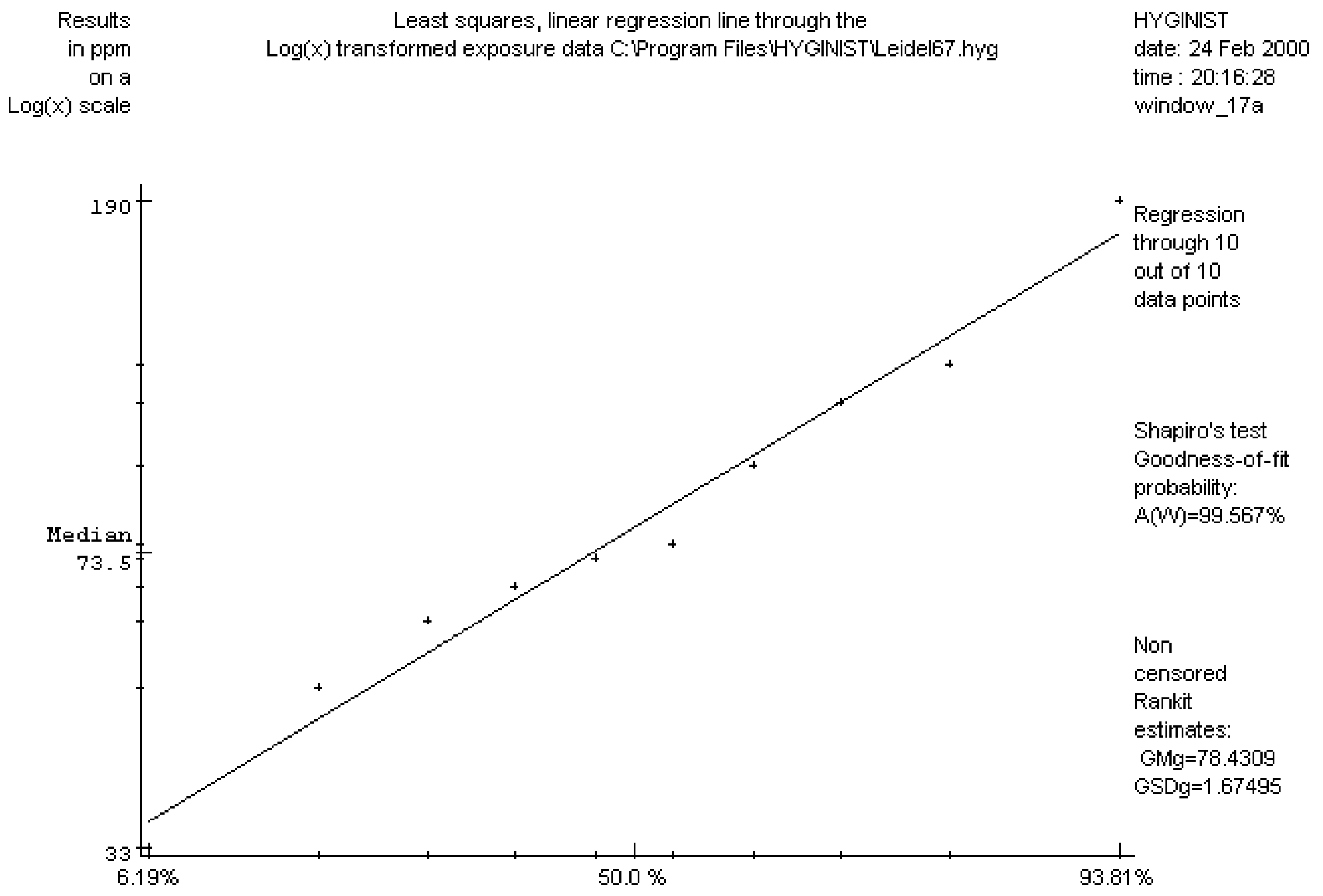
Contents - Index

Figure 4 15 grab sample airborne chlorine concentrations (OWEN716)



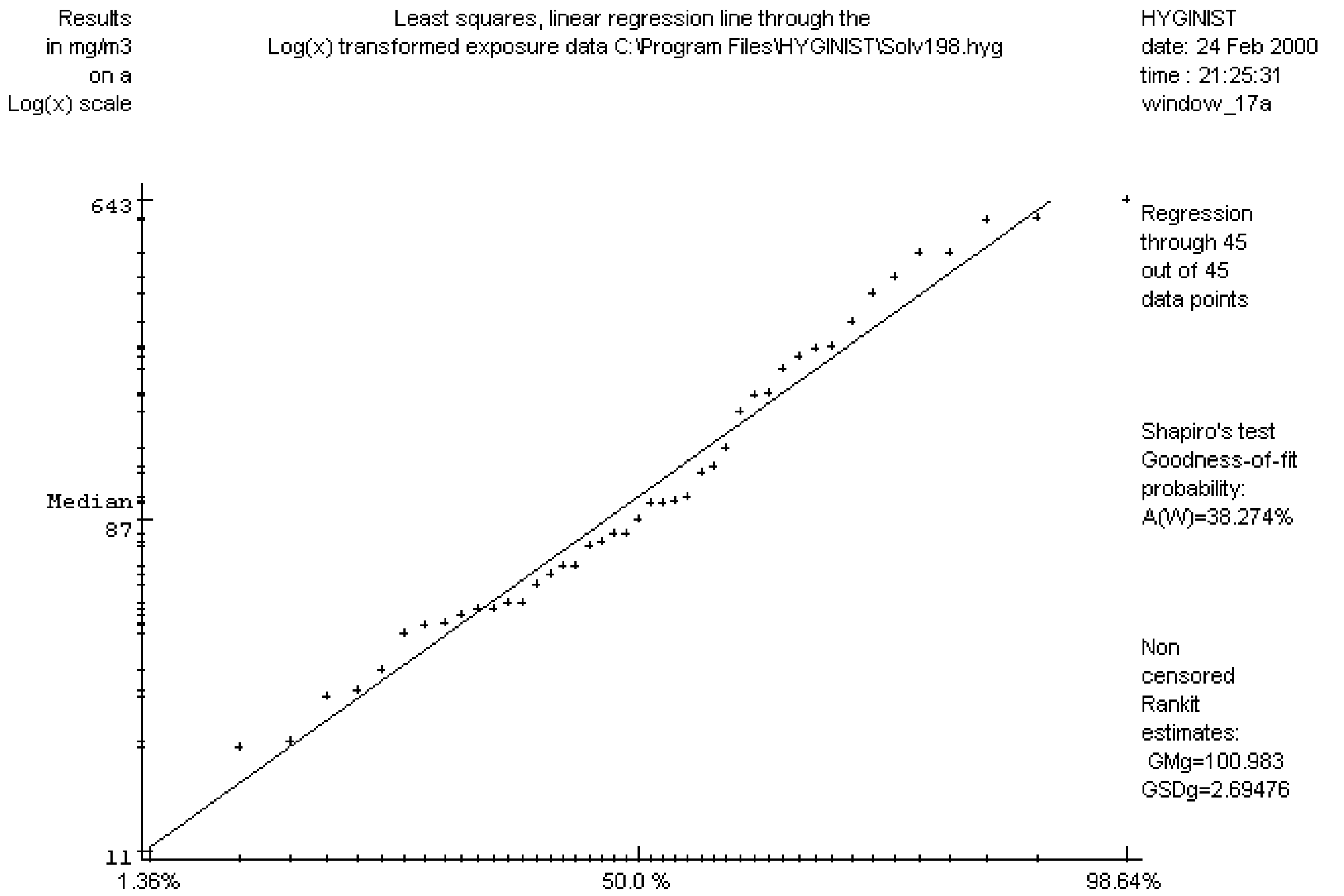
Contents - Index

Figure 1 Ten TWA8 hours Dioxane (LEIDEL.67)



Contents - Index

Figure 2 45 TWA8 hours total hydrocarbons among maintenance painters (Solv198)



Contents - Index

3.1.5 Rankit estimators for Normal descriptive statistics

If you selected the option "untransformed" in Menu Transformation, then the rankit estimators for location and variance (xcg and scg respectively), based on the data pairs in the regression, are displayed.

Table 2 Rankit estimators for Normal descriptive statistics

Example 13 Normal distributions

Example 14 Comparing xcg and scg with estimators from literature

Figure 7 10 Chlorine TWA8 hour. Logarithmic transformed and uncensored (HAW104)

Figure 8 10 Chlorine TWA8 hour. Untransformed and censored (HAW104)

Contents - Index

Table 2 Rankit estimators for Normal descriptive statistics

Estimator Description

M' the number of results between
Formula 3-2 the accuracy limit

$$M' = ul - ll$$

x_{cg} arithmetic mean of the
Formula 3-3 untransformed data, estimated
using the data pairs between
the accuracy limits

$$s_c^g = \frac{\sum_{j=ll}^{j=ul} R_j * c_j - \frac{\sum_{j=ll}^{j=ul} R_j \sum_{j=ll}^{j=ul} c_j}{M'}}{\sum_{j=ll}^{j=ul} R_j^2 - \frac{\left(\sum_{j=ll}^{j=ul} R_j\right)^2}{M'}}$$

s_{cg} standard deviation of the
Formula 3-4 untransformed data, estimated
using the data pairs between
the accuracy limits

$$\bar{x}_c^g = \frac{\sum_{j=ll}^{j=ul} \left[c_j - s_c^g * R_j\right]}{M'}$$

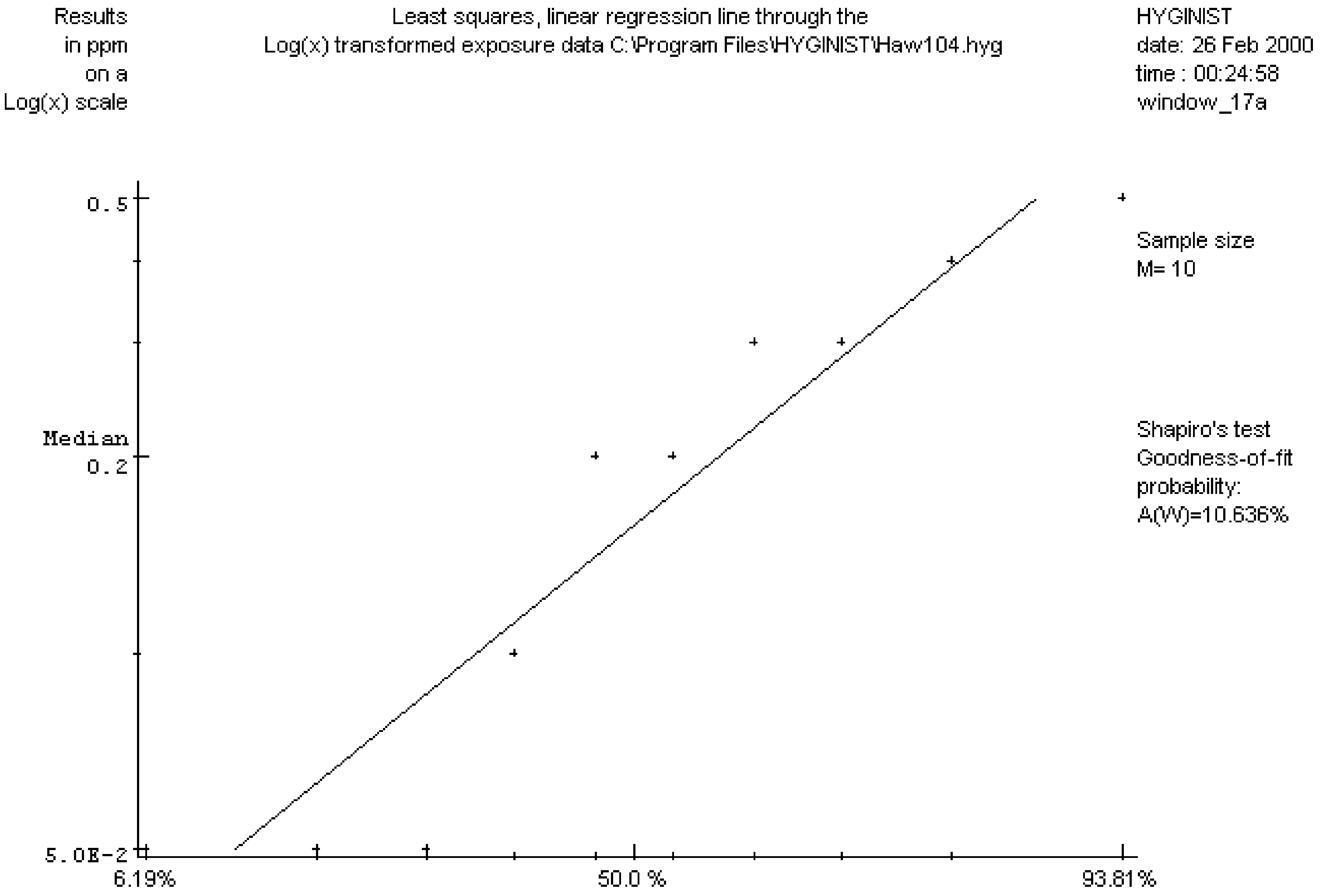
Example 13 Normal distributions

Conformity with other than the Lognormal distributions is found in the following examples from Table 1:

*.HYG file Description $A(W)_{log}$ $A(W)_{unt}$
in % in %
BAR_SI25 Carbon monoxide grab samples 64.4 79.5
HAW117 15 $TWA_{8\ hour}$ total dust (Figure 11, Example 57), 35.6 85.5
DEWELL42 10 Formaldehyde grab samples 6.6 39.9
LEIDEL56 8 $TWA_{20\ min}$ Ethyl Alcohol (Figure 10) $A(W)_{X^2}$=99.2% 41.8 94.9
X07-10 30 concentrations of chemical X. A_{unt} is 17.6 25.9
equivalent to Gustafson (1991 chapter 7)
who finds $A_{unt}(0.9541)$=25.2%,
CONOV195 50 two-digit numbers from telephone book. 2.9 23.1
W_{unt}=.964 is exactly the same as calculated
by Conover (1980 p365)
LEIDEL63 5 Short period exposures Hydrogen sulphide 94.8 95.5
(see Example 57).

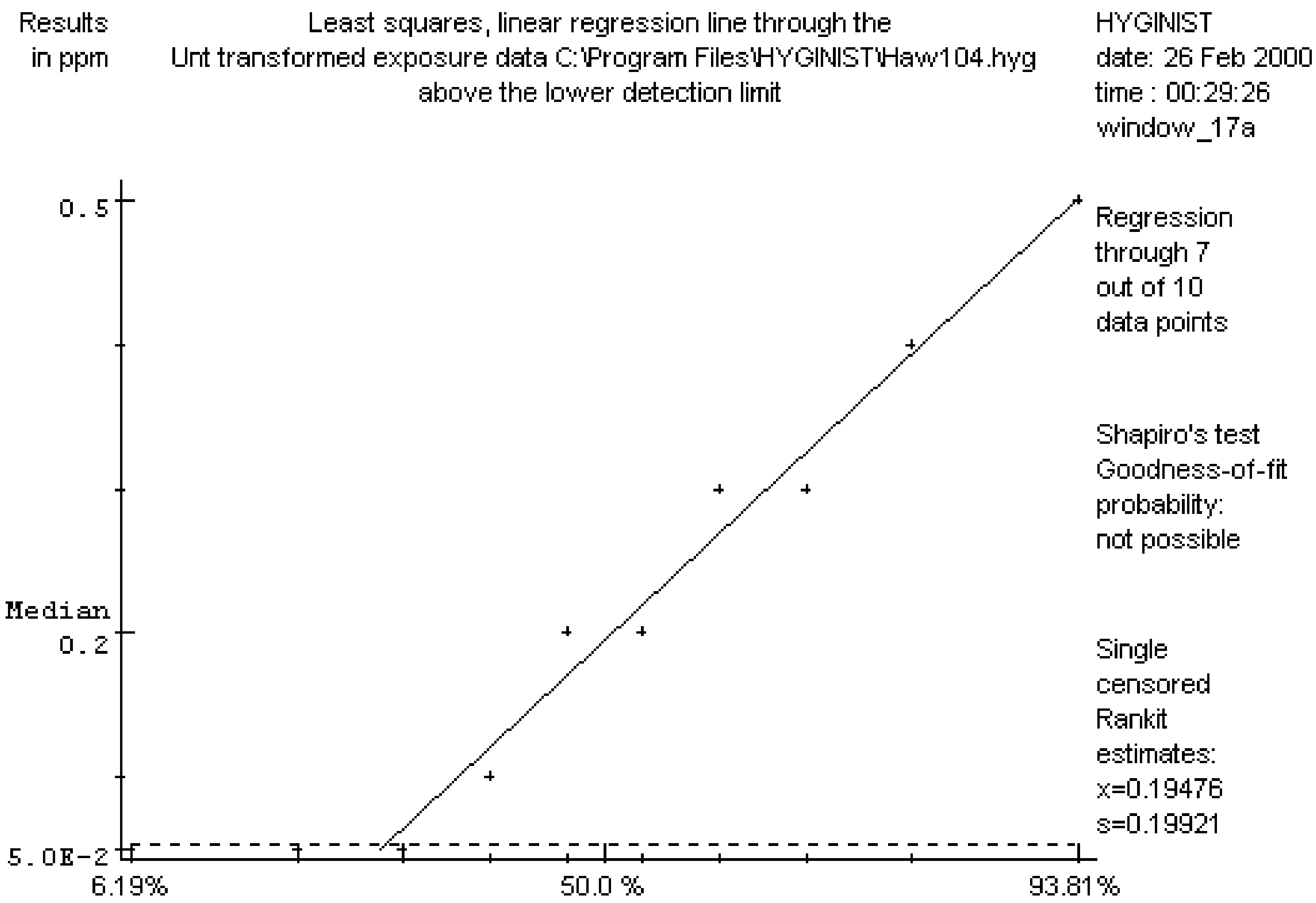
Contents - Index

Figure 7 10 Chlorine TWA8 hour. Logarithmic transformed and uncensored (HAW104)



Contents - Index

Figure 8 10 Chlorine TWA8 hour. Untransformed and censored (HAW104)



Contents - Index

E.3.2 Capture graphics

You can copy bitmaps of all tabs and the 'log-Normal frequency distrbution' using the follwing keys on your keyboard:

- the 'Print Screen' button
- the combination of 'Alt' + 'Print Screen'.

If you then open a picture editor program (like Paint, Photo Editor etc.) or a professional word-processing program, you can paste, edit, select and save the full screen ('Print Screen') or only HYGINIST window ('Alt' + 'Print Screen').

3.2 Tests for normality

The W-test of Shapiro & Wilk (1965) for normality is the most effective one among the existing omnibus tests (see B.1.2). Shapiro (1990 p21), however, warned: "No single [goodness-of-fit] test statistic can give you as much information as a graphical display which shows the extent and types of departures from the hypothesized model".

Contents - Index

3.3 Other transformations

Try different transformations if the Lognormal shows a poor fit. It is not possible to combine two or more of the seven transformations. Different transformations will show similar figures if the variance is small (GSD<1.4).

Menu options "Transformation"

Option Comment
Double Double logarithmic transformation, log{log(ci)}.
If c1<=1, then all results are multiplied with EXP(1)/c1 to keep log(ci) positive
Log The natural logarithm of the result, log(ci), the standard transformation in HYGINIST
Sqrt The square root, %ci
Untranformed Untransformed, ci
Squared Squaring, $(ci)^2$
Exponential The exponential transformation on base e=2.71828.., EXP(ci)=ec.
If c1>44, then all results are first multiplied by 1/c1 to prevent overflow
Reciprocal Reciprocal transformation, 1/ci

Transformations 1 to 6 influence the skewness (asymmetry). The first two decrease the variance and stabilize distributions with high extremes. Transformations 4, 5 and 6 increase variance and stabilize a sample with low extremes. All transformations except the double logarithmic are independent of the location. In Table 3 some distributions and transformations are related.

Table 3 Type of distribution according to the goodness-of-fit with a transformation
Best transformation fit Distribution type
C_i Normal, Gaussian
log(C_i) If s/x ≈ constant for every C_i, then Lognormal
Sqrt(C_i) If $s^2 \approx x$ for every C_i, then the Poisson distribution
1/C_i If s ≈ x^2 for every C_i, then inverse normal

Figure 9 5 TWA8 hour MDI during indoor spraying. Figure 10 8 TWA20 min Ethyl Alcohol.
LogLog transformed (BILAN304) One worker, one shift, untransformed (LEIDEL56)

Example 15 Other transformations

Contents - Index

Figure 9 5 TWA8 hour MDI during indoor spraying. LogLog transformed (BILAN304)

Figure 10 8 TWA20 min Ethyl Alcohol. One worker, one shift (LEIDEL56)

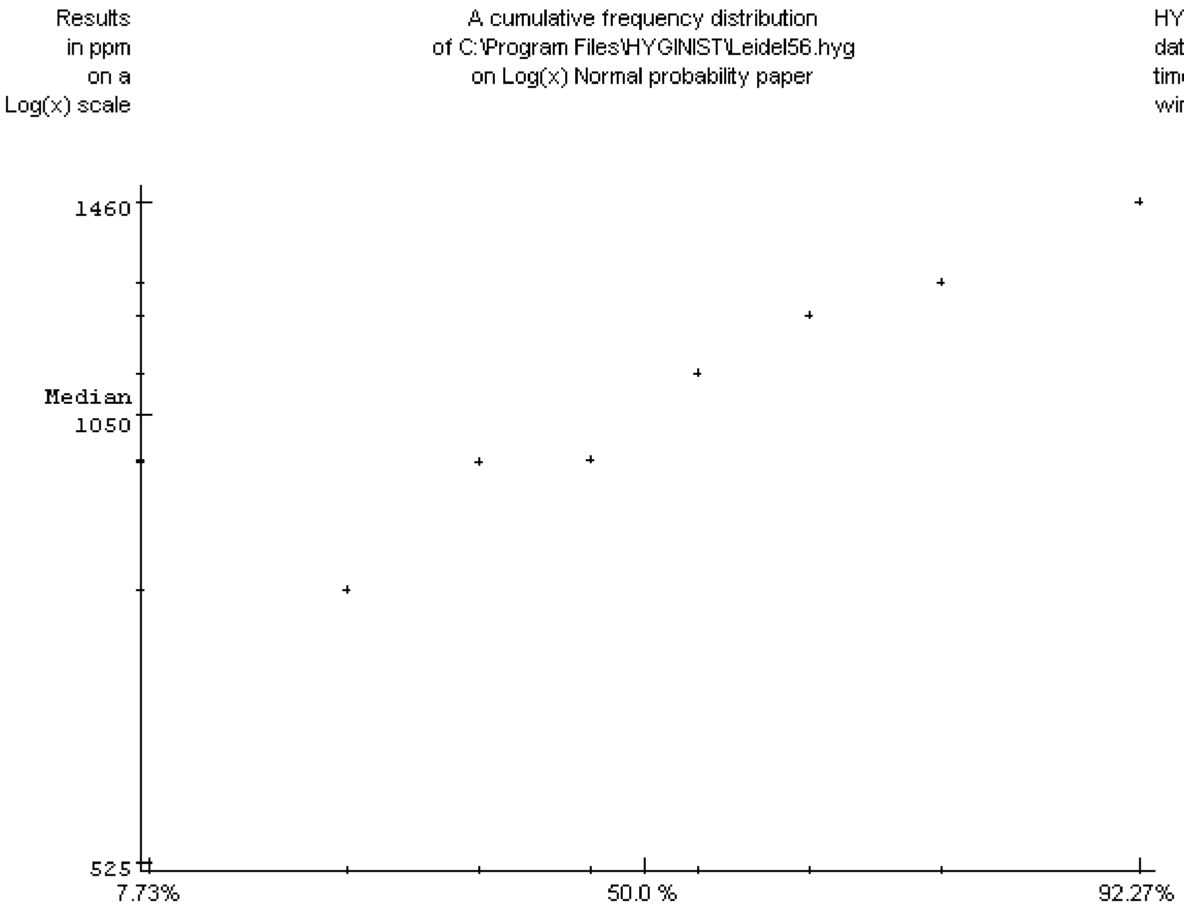


Figure 10 8 TWA20 min Ethyl Alcohol. One worker, one shift (LEIDEL56)

Contents - Index

2.1.1 Range of sample sizes

The sample size M is restricted to integer values between 2 and 2000. Two unequal and uncensored data are at least necessary to calculate a sample variance. The maximum of 2000 arises from the accuracy in calculating rankits (see B.1.1). Also with >2000 data the estimators GM and GSD differ hardly from their descriptive statistics EXP(μ) and EXP(σ). Three or more measurements makes it possible to examine the shape of the sample distribution. More on sample size see 2.2.1.1

Contents - Index

Example 1 Reanalysing exposure data by adapting the number of degrees of freedom

*.HYG file Description
HAW117 To make a file of the 15 $TWA_{8\ hour}$ total dust measurement results from Hawkins (1991 page 117), enter the sample size M=15.
SCHNE224 To estimate the survival duration of items under stress, based on only the 15 items (Schneider 1986) who failed before the end of the observation duration (6 TU's), enter M=15 and the rankit estimators GM_g=9.1 and GSD_g=2.12 (see Figure 5)

2.2.1.3 Sampling Duration

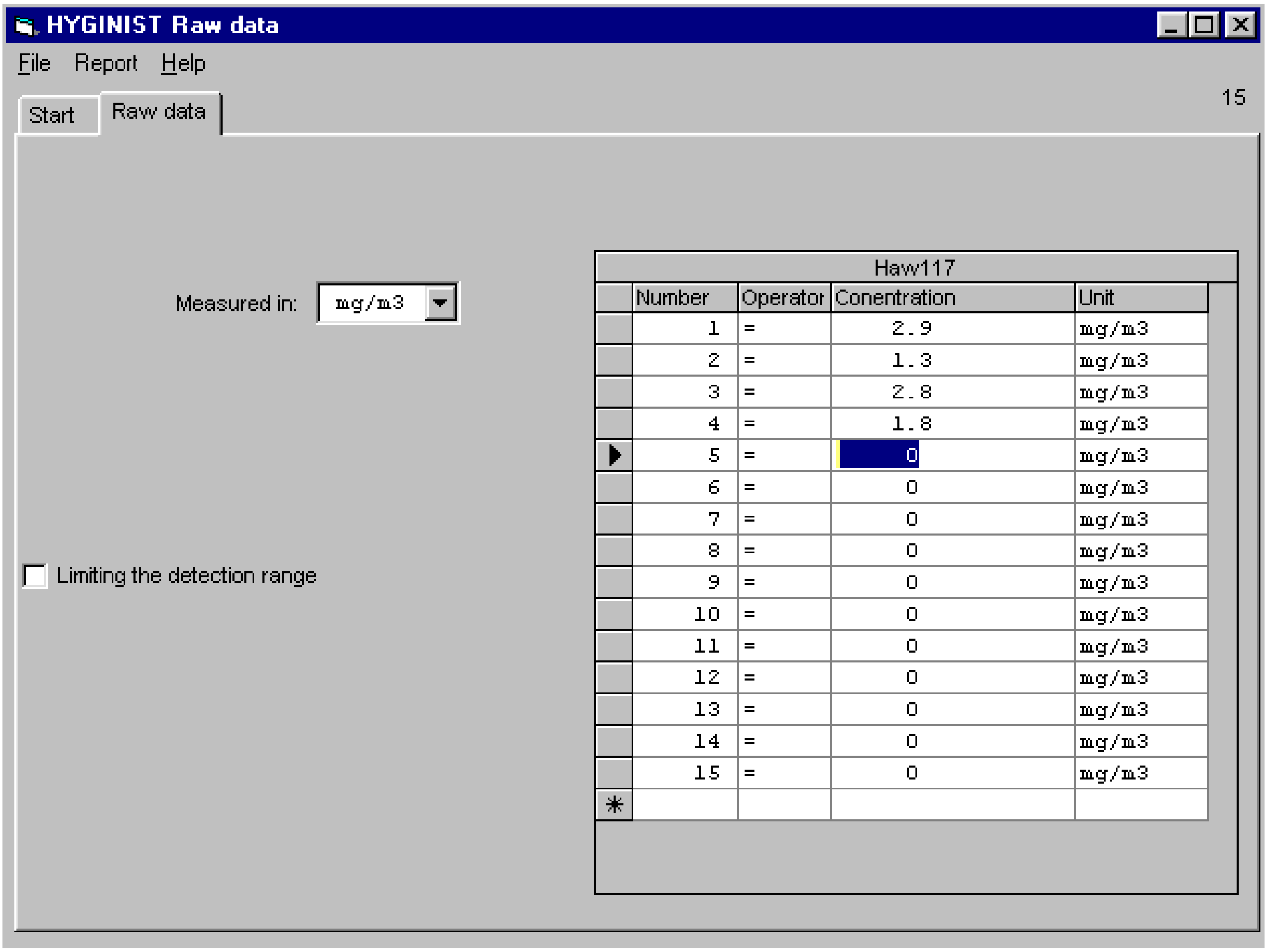
This is the time period the sampling measurements represent:

- a ten seconds sampling period may represent instantaneous sampling,
- a five hours sampling period is often sufficient for a full shift exposure assessment.

Contents - Index

2.2.3.1 Raw data (uncensored)

Fill the cells of the grid with the individual measurement outcome. Use the arrows up and down to move through the grid.



A measurement result is >1D-10 and <1D+10, and consists of: at most 16 digits (double precision) and optional, a floating point and/or an exponent. Enter a positive natural number within the range. The decimal point followed by an E or D, a + of - and one or two digits is recognized as an exponent

After pressing the <Enter> key a result is stored. After the last result is entered the file is sorted (Screen 10, par. 2.3.1). Next transaction is Show and Edit (par. 2.8).

Example 2 Entering raw data

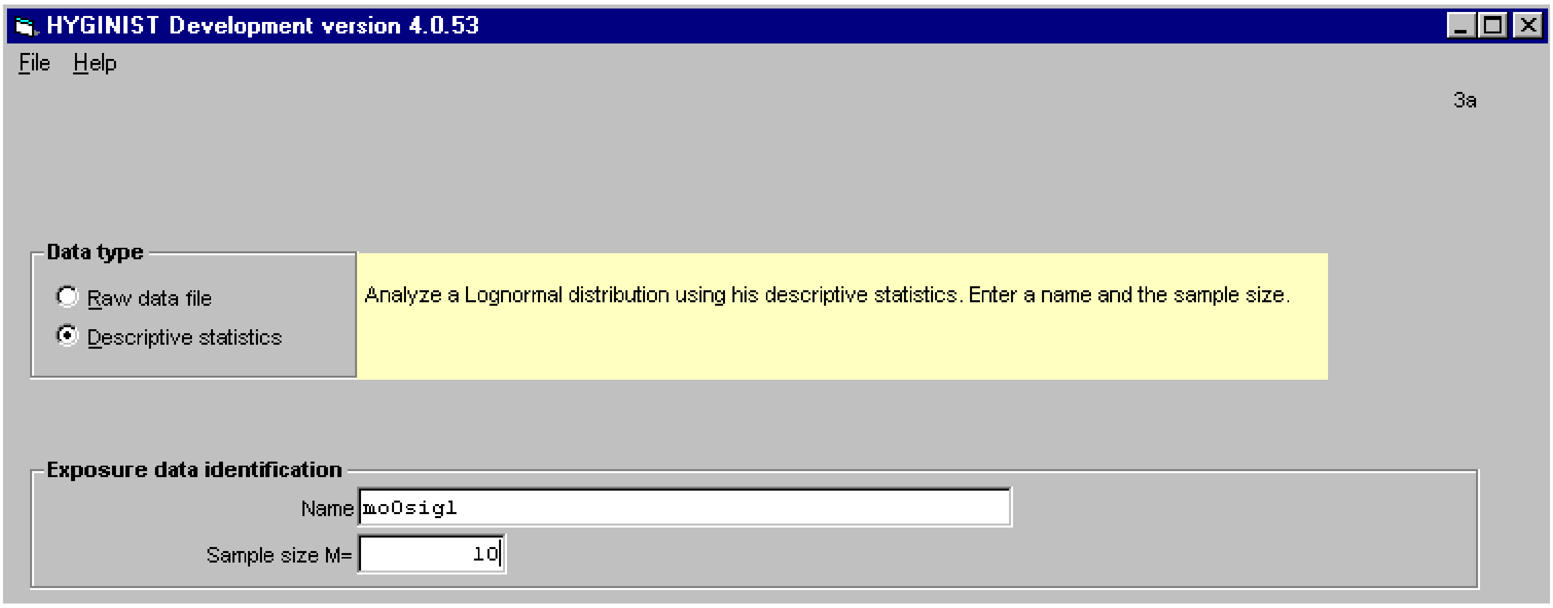
Example 2 Entering raw data

*.HYG file Description
HAW117 Enter the 15 $TWA_{8\ hour}$ random collected total dust samples
(OEL 10 mg/m3) from a similar exposure group
(Hawkins 1991 page 117) in an arbitrary order
2.9, 1.3, 2.8, 1.8, 2.1, 2.5, 2.1, 2.2, 1.9, 2.0, 2.5, 1.8, 2.2, 2.4, 2,5

Contents - Index

2.2.5 Descriptive statistics

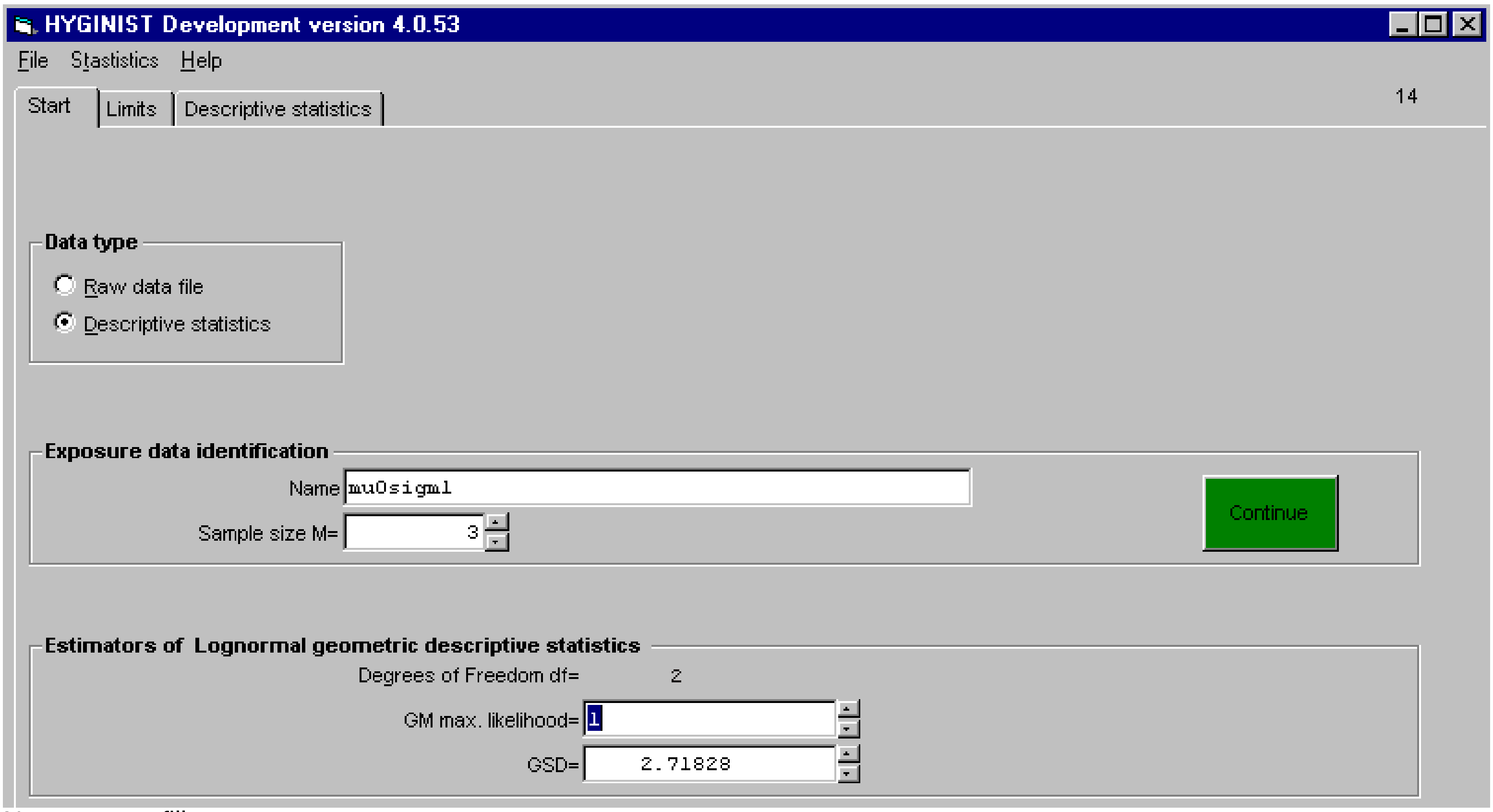
If you click the option button 'Descriptive statistics' in frame 'Data type' then the frame 'Exposure data identification' appears:



In the frame 'Exposure data identification' you can fill out :

- a name (obligate)
- A sample size (obligate)

If name and sample size are entered then the frame 'Estimators of Lognormal geometric descriptive statitics' appears.



Now you can fill out:

- a Geometric mean value between 0.0001≤ GM ≤ 10000

- a Geometric standard deviation value between 1.001≤GSD≤1000

Entering estimators of the Lognormal descriptive statistics:
- assumes the exposure data originate from a Lognormal distribution,
- prohibits the evaluation of shape (chapter 3)

Example 4 Enter GM and GSD

Form more information on GM and GSD see B.2

Contents - Index

Example 4 Enter GM and GSD

*.HYG file Description

- To demonstrate "minimise sample size" in routine long-term compliance control, GSD=EXP(1)=2.71828 and GM=1 are entered using the file name MU0SIG1.HYG.

OWEN716 Reanalyse the airborne Chlorine $TWA_{8\ hour}$ (Owen 1980 p716) using only the 8 uncensored grab sample data within the accuracy range of .25 and 9 PPM. Enter (See Figure 4) in Screen 3a M=8 and in Screen 7ab the estimators GM_g=.88 PPM and GSD_g=6.81 (Continued in Example 10).

Contents - Index

2.1 Sample requirements

The numerical requirements on exposure values prevent random errors and the sample range (in numbers and numerical values) from influencing the validity and exceeding the range of accuracy.
Requirements are given for:

- sample size (2.1.1)
- measurement outcome (2.1.2)
- Rounding and grouping interval (2.1.3)

2.1.2 Range of valid exposure data outcome

Sample measurement values are restricted:
- between c(1) >=0.0000000001 and c(M) <=10000000000.
- in the ratio of their extremes: 1.0001<c(M)/c(1) < 10000000000.

Sample measurements and the Lognormal descriptive statistics are further restricted in the range of the:
- geometric mean 0.0000000001≤GM ≤ 10000000000,
- geometric standard deviation 1<GSD ≤ 10000.

A GSD=1 indicate that all exposure data are equal. The presence of variance within a series measurements is a requirement for performing statistical analysis

Contents - Index

2.1.3 Rounding and grouping interval

The measurement method or the coarse rounding of measurement outcome may overshadow the workplace variance and thus influence the shape of sample distribution and the validity of the evaluation. To prevent this the arithmetic standard deviation w should be much larger than the rounding and grouping interval ∇C.

Contents - Index

2.2 Data entry

Choose between descriptive statistics or raw data (2.2.3). In that way it is possible to analyze:

- new series of measurement results,
- series of measurement results that are exported from exposure databases,
- literature or historical data from which only sample size and the estimators GM and GSD of the Lognormal descriptive statistics are reported.

You can enter exposure data :

- by keyboard (2.2.1),
- from a file in a directory (2.2.4).

Contents - Index

2.2.1 Keyboard

Entering exposure data involves:

- sample size (2.2.1.1),
- the units of measurement (2.2.1.2),
- the sampling duration of measurement (2.2.1.3),
- characterization of the exposure by a name (2.2.2),
- the choice between raw data and descriptive statistics entry (2.2.3).

Contents - Index

2.9 *.HYG file editors

Many software programs can create, read, write, expand, compress and edit *.HYG files. For example spread sheet Programs like Lotus or Excel, database programs like Access and dBase , Text editors like Notepad, Editpad, wordprocessing programs like Word and WordPerfect and many others (e.g. SPSSPC, Basic, C). They can, like HYGINIST, process one column (field), numerical SDF or DELIMITED files with the extension *.HYG.

Example 6 Making a *.HYG file with an other program

Example 7 Generating a Lognormal Monte Carlo *.HYG file using Basic

Example 8 Making a *.HYG file using dBase

Contents - Index

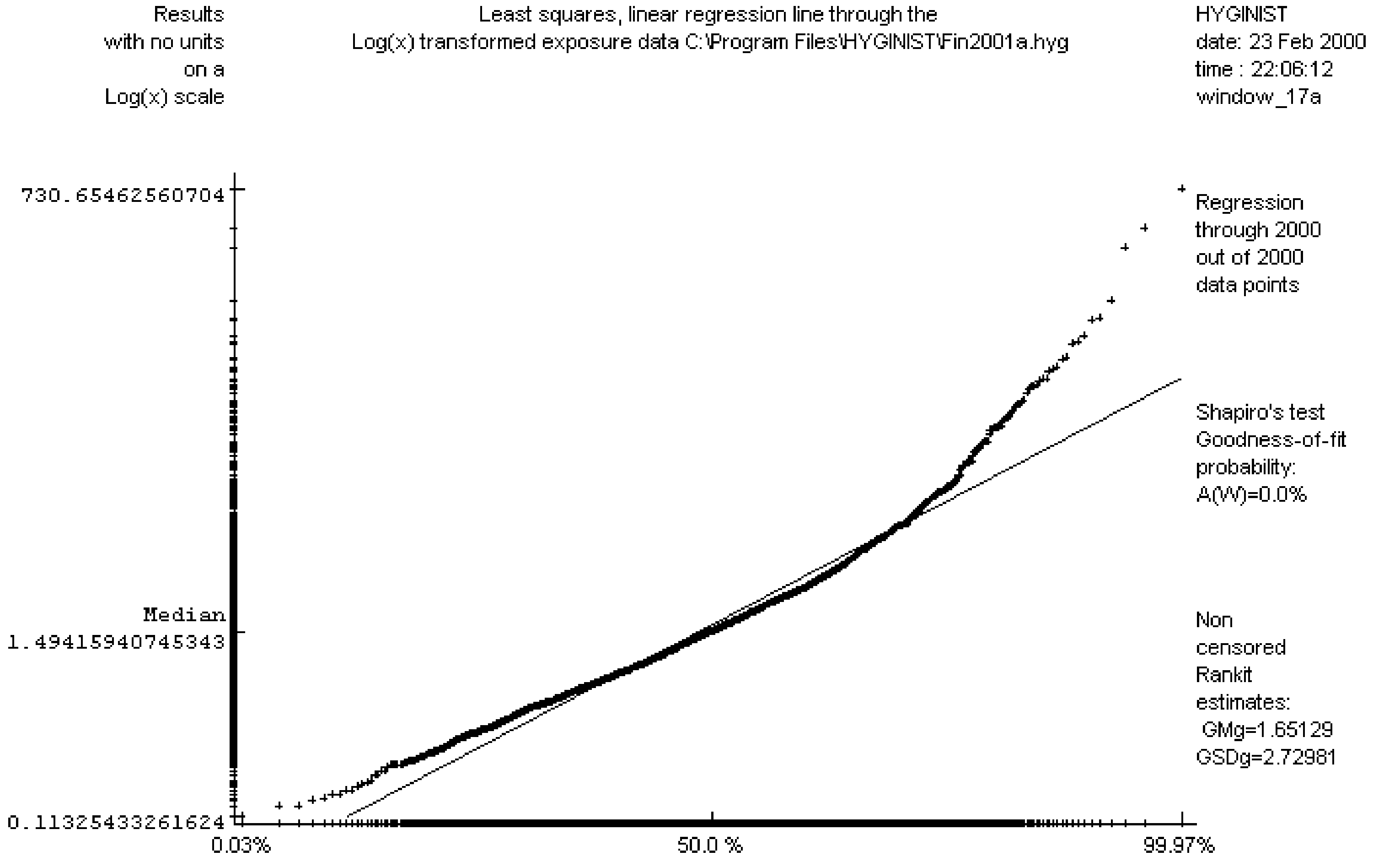
Example 6 Making a *.HYG file with an other program

HAW104.HYG is made (<.05, <.05, <.05, 0.1, 0.2, 0.2, 0.2, 0.3, 0.3, 0.4,0.5) using Word.
The following numbers are typed on an empty page:
.05,.05,.05,.1,.2,.2,.2,.3,.3,.4,.5
The file is saved by clicking the mouse on menu button <File>, <Save as>
Now choose option <text file> and type HAW104.HYG<Enter>
More convenient is entering the data in a column using the <Enter> key and save it as an text file.
.05
.05
.05
.1
.2
.2
.2
.3
.3
.4
.5

Contents - Index

Example 7 Generating a Lognormal Monte Carlo *.HYG file using Basic

The file HALD151 with 500 rivet head diameters is easily made using dBase because outcome were allocated in only 9 grouping intervals (see Figure 24



).

The dBase control centre was left using F10 and 'Exit to the dot prompt'. The following command lines were used:

- Command CREATE <database name>
- Field Name (W10 characters),
- Field type <Numeric>,
- Width 10 digits (including the decimal point and sign),
- Dec   5 decimal places,
- Index   N,

Press <Ctrl> + <End> to stop CREATE database.
Command a MODIFY COMMAND <program name> and program the following lines:
DO WHILE RECNO()<=500
APPEND BLANK
ENDDO
Press <Ctrl> + <End> to stop MODIFY COMMAND,

Command DO <program name>.
Now command the following lines from the dBase prompt
REPLACE <Field name> WITH 13.39 FOR RECNO()=>1   .AND. RECNO()<191

```
REPLACE <Field name> WITH 13.42 FOR RECNO()=>191 .AND. RECNO()<287
REPLACE <Field name> WITH 13.47 FOR RECNO()=>287 .AND. RECNO()<359
REPLACE <Field name> WITH 13.52 FOR RECNO()=>359 .AND. RECNO()<427
REPLACE <Field name> WITH 13.57 FOR RECNO()=>427 .AND. RECNO()<468
REPLACE <Field name> WITH 13.62 FOR RECNO()=>468 .AND. RECNO()<486
REPLACE <Field name> WITH 13.67 FOR RECNO()=>486 .AND. RECNO()<498
REPLACE <Field name> WITH 13.72 FOR RECNO()=>498 .AND. RECNO()<500
REPLACE <Field name> WITH 13.77 FOR RECNO()=500
```

for entering the measurement results in the *.DBF file. Finally command a

```
COPY TO HALD151.HYG FIELD <Field name> DELIMITED
```

for making the HALD151.HYG file.

Contents - Index

Example 8 Making a *.HYG file using dBase

With the algorithm of Abramowitz (1971, p 953, direct method) the 2000 deviates of the standard Lognormal distribution were generated and stored in the one-field ASCII delimited file (RANDOM20.HYG) using the following BASIC subroutine:

```
DIM C(2000)
L% = 0
RANDOMIZE CDBL(TIMER)
FOR I% = 1 TO 2000
 IF L% = 1 THEN
  Z = EXP(Y2 * V): L% = 0
 ELSE
  DO
   DO
    Y1 = 2# * CDBL(RND) - 1#
    Y2 = 2# * CDBL(RND) - 1#
    W = Y1 ^ 2 + Y2 ^ 2
   LOOP WHILE W >= 1#
   V = SQR(-2 * LOG(W) / W)
   Z = EXP(Y1 * V): L% = 1
  LOOP WHILE Z <= 1D+300# AND Z >= 1D-300#
 ENDIF
 C(I%) = Z
NEXT I%
```

Contents - Index

Getting Started (Summary)

A statistical tool like HYGINIST is useful if the effectiveness of working condition control measures is assessed using discontinuous exposure measurements.

HYGINIST includes 6 transactions:
1. data entry and mutation,
2. examining distribution shape,
3. estimating the descriptive statistics,
4. extrapolation to unsampled periods and compliance test against a limit value,
5. comparing the exposure data with other sets of descriptive statistics,
6. establishing the minimum sample size for an unbiased estimate of long-term compliance control.

This Help further explains:

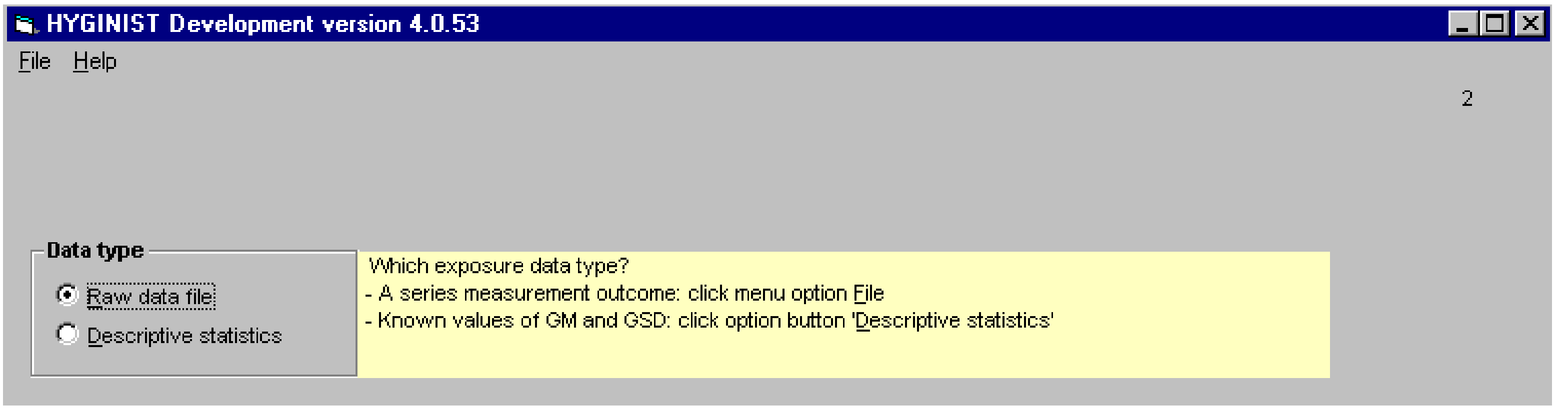
- How to apply the HYGINIST statistical methods in working conditions control
- The many aspects of the Lognormal distribution
- Installing and running the program
- Program operations
- Register for regular use.

.

If HYGINIST is started, then the following Splash screen appears.
After installation it will include the next Login box.

See Annex C.5.1. on how to fill out this form is

If the Login form is filled out correctly then the HYGINIST start page appears after a few seconds, in which all form and constants are loaded. Now you can start with the exposure data management and analysis.



I hope this program support your needs.
Please inform me by mail ihpc@planet.nl on your experience and don't forget checking regularly on the HYGINIST homepage for updates of the program and the helpfile.

Regards
Theo Scheffers

Click here for a summary on the Goodness-of-fit methods used.
Click here for a summary on the Extrapolation methods used.

4 Sample characteristics

Tabs 3 ("Limits") and 4 ("Descriptives") makes it possible to:

- overview all sample descriptive statistics (see 4.1)
- change the environmental factors (see 4.2)

Contents - Index

4.1 Descriptive statistics

This tab gives you an overview of different descriptive statistics. It is assumed the exposure data are sampled from a population that can be described using Lognormal descriptive statistics. All calculations are done using the sample estimators GM and GSD of Lognormal descriptive statistics EXP(μ) and EXP(σ) respectively, and the number of degrees of freedom df. Window 20abc shows data from:

- a complete sample (4.1.1),
- a censored sample (4.1.2),
- the entered values for M, GM and GSD (4.1.3)

Contents - Index

4.1.1 Complete sample

This tab gives you an overview of different descriptive statistics calculated from GM, GSD and sample size M.

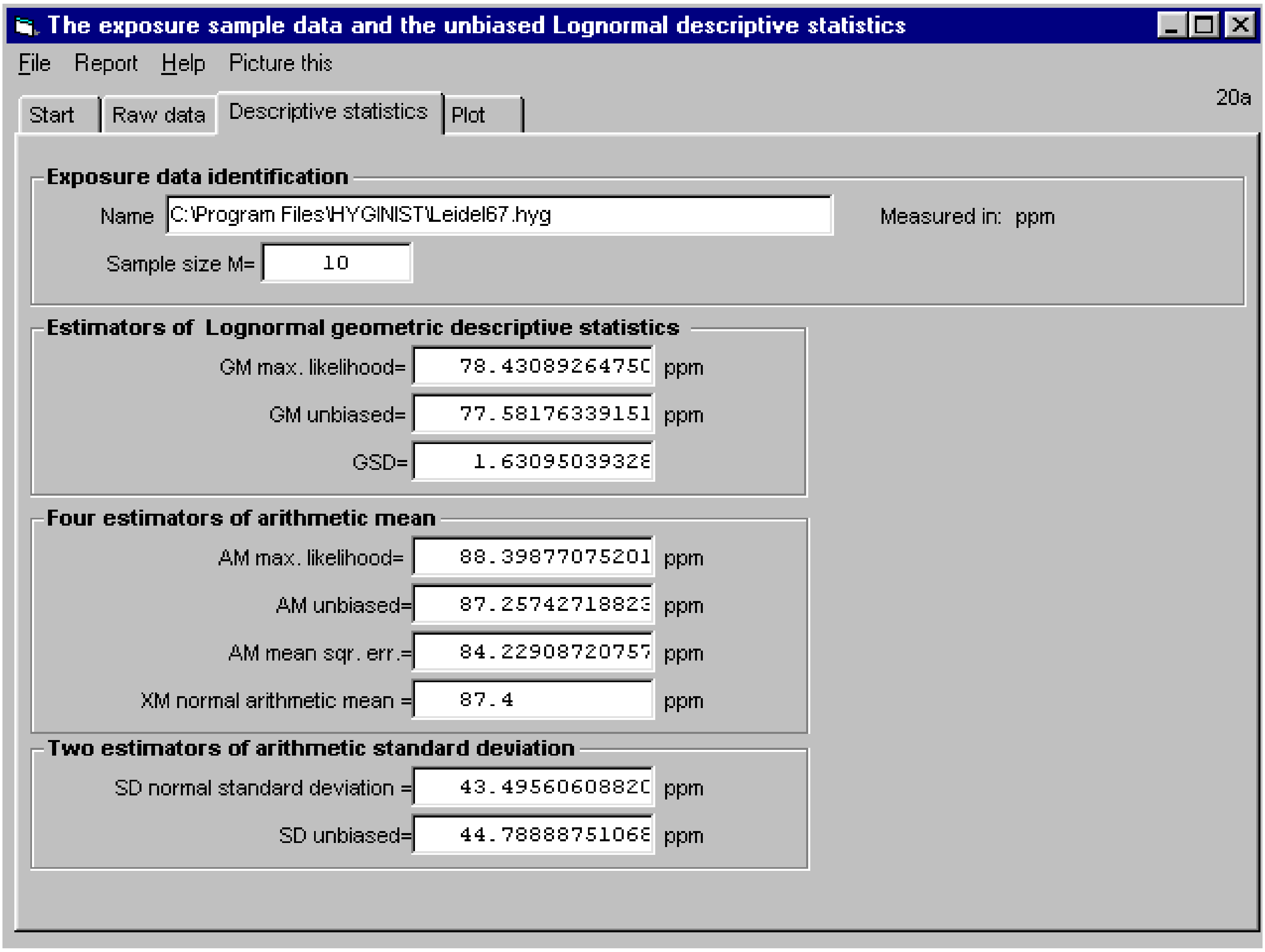


Table 5 Characteristics of the complete sample estimators of the Lognormal descriptive statistics

Example 18 Comparing samples with literature data

Contents - Index

Table 5 Characteristics of the complete sample estimators of the Lognormal descriptive statistics

Estimator Description and properties

df The number of degrees of freedom $df = M - 1$
Formula 4-1

GM Estimator of the geometric mean EXP(μ). $GM = EXP\left(\bar{x}\right)$
Formula 4-2 Measure of source strength (Seixas 1988, Rock 1982)

$$\bar{x} = \frac{\sum_{i=1}^{M} x_i}{M}$$

0 < GM < ∞
Dimension c_i

If GM > H, then at least one c_i > H. $x_i = \log(c_i)$
Biased estimator of EXP(μ): If M≤10 and GSD>2.7, then GM overestimated at least 5% on average.

GM_U Like GM (see above) with the following exceptions:
Formula 4-3 Unbiased estimator of EXP(μ)

$$GM^u = GM * \phi\left(-\frac{s^2}{4}\right)$$

with

$$\Phi(t) = \sum_{k=0}^{\infty}\left(\frac{\Gamma_M}{\Gamma_{k,M}} * \frac{t^k}{k!}\right)$$

GM_U<GM for all values of M and s
If GSD/M_2>6, then "Cannot calculate unbiased estimator GMu!" is displayed.
Relevant for the noise level pressure kPa.

GSD Estimator of the geometric standard deviation EXP(σ)
Formula 4-4 Measure of scale in a Lognormal distribution

$GSD = EXP(s)$
No dimension
>1

$$s = \sqrt{\frac{\sum_{i=1}^{M} \left(x_i - \bar{x}\right)^2}{M - 1}}$$

Unbiased estimator with minimum variance.
Normal values in the workplace atmosphere
ranges between 1.4 and 3.5.

AM Estimator of the arithmetic mean β
Formula 4-5 Calculated using the following formula.

$$AM = GM * \phi\left(\frac{s^2}{2}\right)$$

with

$$\Phi(t) = 1 + \frac{(M-1)*t}{M} + \frac{(M-1)^3 * t^2}{M^2 * (M+1) * 2!} + \frac{(M-1)^5 * t^3}{M^3 * (M+1) * (M+3) * 3!} + \ldots$$

Unbiased minimum variance estimator
$TWA_{reference\ period}$ from series short period measurements
within a reference period (Bar-Shalom 1975)
The mean dose of a cumulating agent (Seixas 1988)
from series of more days $TWA_{8\ hour}$
Dimension $c_i > 0$

w^2 Unbiased minimum variance estimator
Formula 4-6 of the arithmetic variance w^2 (the 2nd moment)

$$w^2 = GM^2 * \left[\phi\left(\frac{s^2}{2}\right) - \phi\left(\frac{M-2}{M-1} * s^2\right)\right]$$

Dimension c_i^2
>0
Inverse proportional with ventilation rate
and the square of the volume (Roach 1977 p67).

w/AM The sample coefficient of variation, to be compared with CV_t,
rounding error and grouping interval ratio -C/AM.

Example 18 Comparing samples with literature data

*.HYG file Description
LEIDL104 24 TWA8 hour Methyl Methacrylate results in GM=34.5 PPM and GSD= 1.89
LEIDEL67 10 TWA8 hour Dioxane (see Figure 1) result in GM=78.4 PPM and GSD= 1.63.
CHIP123 10 TWA8 hour lead result in GM=2.19 and GSD=1.78.

The first two series provide estimators of the descriptive statistics which are equal to the literature values (Leidel 1977 pages 67 and 104). Booher (1988) calculated GM=2.36 and GSD=1.76 from the lead exposure data among chippers. The difference cannot be explained.

4.1.2 Censored sample

This tab gives you an overview of different descriptive statistics calculated from the regression estimators of GM and GSD.

4.1.3 Entered values of GM and GSD

Not yet filled

4.2 Environmental factors

Not yet filled

6 COMPARING EXPOSURE DATA

The fifth transaction compares the exposure estimators GM and GSD with:

- estimators from a second sample,
- true population descriptive statistics,

Valid reference data are the true Lognormal descriptive statistics EXP(μ) and EXP(σ), or the sample estimators GM and GSD and the sample size.
To compare more than two data sets, apply ANalysis Of VAriance on log-transformed exposure data (Snedecor 1980). ANOVA techniques are included in standard software like Microsoft® EXCEL.

The current developments in occupational health and safety care makes that:
- more exposure data are collected in industrial hygiene (surveys, routine programs etc.),
- (one or multi) year Health and Safety plans and engineering directives, plant lay-out, control measures ask for more quantified exposure information,
- dose assessment becomes more important in occupational epidemiology.

This leads to an increasing need to collect, analyze and compare exposure data, not against limit H, but against other location and scale data:
- in time (changing control measures, in longitudinal epidemiology),
- between workplaces (different technical, organizational or behaviors control measures),
- between exposure groups (differences in dose in epidemiology),
- within exposure groups ({dis}homogeneity, identifying worst cases, time trends),
- against true population values ("gold" standards).

This chapter explains with a lot of examples how this is done. For more information see chapter 6.1

Contents - Index

6.2 Two sample difference

Sample size M_2 (see 6.2.1) and the estimators GSD_2 (see 6.2.2.) and GM_2 (see 6.2.3) of a second sample should be entered. It should be established first, however, that these data are from a Lognormal distribution. Use the raw data and the goodness-of-fit transaction described in chapter 3.

6.2.1 Reference sample size

The maximum size of the reference sample (M_2=30000) is more than is permitted in data entry from disk or by keyboard (see 2.2.1.1).

Contents - Index

6.2.2 Difference in scale

The statistic for testing if both GSD`s originate from the population is (Snedecor 1980):

$$F_{df,df_2} = \left(\frac{\mathrm{LOG(GSD)}}{\mathrm{LOG(GSD_2)}}\right)^2 * \frac{\mathrm{df}_2}{\mathrm{df}}$$ Formule 6-1

Under the null hypothesis follows F the variance ratio or Fisher distribution (Abramowitz 1970 26.6):

$$A_{F|df_1,df2} = \frac{df_1^{\left(\frac{df_2}{2}\right)} * df_2^{\left(\frac{df_1}{2}\right)} * \Gamma\left(\frac{df_1+df_2}{2}\right)}{\Gamma\left(\frac{df_1}{2}\right) * \Gamma\left(\frac{df_2}{2}\right)} * \int_{\frac{1}{F}}^{F} x^{\left(\frac{df_1-2}{2}\right)} * (df_2 + df_1 * x)^{\left(\frac{df_1+df_2}{2}\right)} dx$$

$$\text{with } F \geq 0$$ Formule 6-2

F df1 df2 $A_{GSD=GSD2}=100*A_{F|df,df2}$ %
W10-31 <∞ 0 %
D10+31 <∞ 0 %
ABS(F-1)W10-10 df1=df2 100 %
other values odd odd Algorithm 26.6.8
other values even odd or (even & $df_1>df_2$) Algorithm 26.6.4
other values odd or (even & df_1Wdf_2) even Algorithm 26.6.5
other values df_1*df_2D32768 Normal approximate 26.6.15

All algorithms are from Abramowitz (1970)

Example 39

Contents - Index

6.3 Differences with population descriptive statistics

Reasons to compare sample estimates with true values of EXP(μ) and EXP(σ) are:

- the sample size van reference data are large or unknown (e.g M2D100),
- testing against program goals or golden standards in industrial hygiene,
- finding the confidence range for EXP(μ) and EXP(σ), using trial and error.

Contents - Index

6.1 Sample or population reference values

This transaction establishes differences and trends in exposure data when at least one of the following sets of reference values are known:

- GSD_2 and size M_2 of a second sample (see 6.2),
- EXP(σ) (see 6.3.1),
- EXP(μ) (see 6.3.2).

To compare GM's it is important to know if GSD's:
- are equal (see 6.2.3.1), or
- differ significantly (see 6.2.3.2).

Combining the exposure data from two samples is described in 6.2.4. To compare three or more data series see 6.4. To append the worksheet in the <file name>.LOG file (for the report function see) press <file> and <print> after the calculations took place.

Location and scale are compared using the classical parametric Student methods and formulas (see Snedecor 1980, chapters 5 and 6). The subjoined table 12 combines names found in literature for inference statistics, and displays the used names (in bold) and symbols.

Contents - Index

6.3.1 EXP(σ)

The test statistic that compares log(GSD)² with σ² is (Snedecor 1980 5.11):

$$\chi^2 = df * \left(\frac{\text{LOG(GSD)}}{\sigma}\right)^2$$ Formule A-8

The test statistic χ^2 follows, under the null hypothesis, a Chi-square distribution with a number of degrees of freedom df=M-1 (df=(M+M`)\2-1 or df=2M` for censored samples). For the theory on the confidence interval of GSD, see Land (1988 page 98-99).

$$Q_{\chi^2,df} = \frac{\int_{\chi^2}^{\infty} t^{\left(\frac{df}{2}-1\right)} e^{-\frac{t}{2}}\, dt}{2^{\frac{df}{2}} * \Gamma\left(\frac{df}{2}\right)}$$ Formule A-9

Calculating the chi-square probability distribution
ABS(χ) Df $A_{GSD=\sigma}$
all values <∞ if $Q_{\chi^2,df}<0.5$, then $A_{GSD=\sigma}=200*Q_{\chi^2,df}$ % else $A_{GSD=\sigma}=200*(1-Q_{\chi^2,df})$ %
$<10_{-16}$ <∞ 0 %
W24 and W200 and even Algorithm 26.4.4 (Abramowitz 1970)
W24 and W200 and odd Algorithm 26.4.5 (Abramowitz 1970)
>24 or >200 Normal approximate Algorithm 26.4.14 (Abramowitz 1970)

Contents - Index

6.3.2 EXP(μ)

The test statistic that compares log(GM) with μ is (Snedecor 1980, 5.7):

$$t_{df} = \sqrt{df+1} * \frac{\mu - \mathrm{LOG(GM)}}{\mathrm{LOG(GSD)}}$$ Formule 6-10

Formula 6.10 follows, under the null hypothesis, the Student distribution. The two-sided probability A% is calculated using tdf and dft from 6.10 and the Student distribution. The theory on the confidence interval of GM is described by Land (1988 page 93-95 3.1.2 and 3.2).

Example 43 Censored sample confidence limits

Contents - Index

Example 43 Censored sample confidence limits

*.HYG file description
GUPTA271 In Example 53 (mice, inoculated with tuberculosis) different estimators of EXP(μ) and EXP(σ) are compared. Using the rankit estimators, df=7 and the confidence factors the two-sided 95% CI is 46.7<EXP(μ)<66.1 and 1.15<EXP(σ)<1.56, which is comparable with Schneider`s (1988 p196 example 6.1) two-sided 95% CI's: 48.1<EXP(μ)<66.7 and 1.13<EXP(σ)<1.59.
HAW110 Hawkins (1991 Appendix III p133) uses defective formulas for the upper (GUCL) and lower (GLCL) "confidence limits on the true geometric mean" (AM_{ml} in stead of GM based). For total dust (GM=2.6 mg/m3, GSD=1.77) he calculates two-sided 95% confidence interval for EXP(μ), of G_{LCL}=2.28 and G_{UCL}=4.1 mg/m3. A more appropriate interval, using the inverse of formula 6.10, is 1.94-3.49 mg/m3.

6.2.3.1 Comparing GMs with matching GSDs

The statistic for testing if both GM`s originate from the population sample base is (Snedecor 1980, 6.9):

$$t_{df_t} = \frac{LOG\left(\frac{GM_2}{GM}\right)}{S_w} * \sqrt{\frac{(df+1)*M_2}{df+1+M_2}}$$ Formule 6-3

The weighted logarithmic standard deviation s_w is calculated using:

$$s_{gew} = \sqrt{\frac{df*s+df_2*s_2}{df_t}}$$

with

$$s = LOG(GSD)^2 \quad s_2 = LOG(GSD_2)^2 \text{ and } df_t = df + df_2$$ Formule 6-4

Formula 6.3 follows, under the null hypothesis, the Student distribution. The two-sided probability A% is calculated using t_{df} and df_t from formula 6.3, and the Student distribution (see formula 5.6).

Example 40

Contents - Index

6.2.4 Combining two samples

If two samples originate from the same population base and:

- the two samples are mutually independent, or
- the two samples were separated from one main sample,

then the number of degrees of freedom and the estimators GM and GSD of both samples can be combined into new estimators GM_t and GSD_t for that population base, using the next logical formulas:

$$GM_t = EXP\left(\frac{(df+1) * LOG(GM) + M_2 * LOG(GM_2)}{M_t}\right)$$

$$\text{with } M_t = df + M_2 + 1$$

Formule 6-6

$$GSD_t = EXP\left(\sqrt{s_t + x_t - \frac{M_t}{M_t - 1} * LOG(GM_t)^2}\right)$$

with

$$s_t = \frac{df * LOG(GSD)^2 + df_2 * LOG(GSD_2)^2}{M_t - 1} \quad x_t = \frac{(df+1) * LOG(GM)^2 + M_2 * LOG(GM_2)^2}{M_t}$$

Formule 6-7

If the two samples are mutually independent, then GSDs and GMs may not differ significantly. If the samples were separated from one main sample, then individual raw data may not be included in both samples.

Example 42 Comparing two GSDs

Contents - Index

6.4 Comparing three or more samples

To compare three or more data series, use:

- parametric analysis of variance (ANOVA) and regression analysis,
- distribution free methods (Snedecor 1980 chapters 6 & 10.9, Siegel 1956 chapter 5) like:
- Wilcoxon (two samples),
- Kruskall-Wallis (non parametric ANOVA, three and more samples),
- Page (1963, trend in three and more samples), and
- Spearman or Kendall (rank correlation).

These methods are available in stat/graph packages like BMDP, GLIM, TRUE EPISTAT, SAS-PC, SPSSPC and BMDP.

Example 44 Regression analysis in exposure control chart

Contents - Index

7. SAMPLE SIZE

The tabs <Upper Limit UTL> and <Mean UCL> may help to calculate the smallest sample size for which a series TWA measurements will estimate compliance ($A_{C>H} \leq \alpha$) or non-compliance ($A_{C>H} \geq 100-\alpha$) using the unbiased method of Wilks, the upper confidence limit .

The sample size M_2 is calculated using:

- the desired percentage U%=100-α,
- the industrial hygiene limit value H ,
- the estimators GM and GSD or GM_g and GSD_g and the proportion results within the accuracy range (screen 20b),
- the inverse algorithm of the standard normal and the unbiased Student .

Values for GM, H and GSD can be derived from:

- calculations on emission and ventilation (A.2.1),
- exposure assessment programs or inventories done before (A.3),
- data from literature.

The standard normal estimator A_Z (formula 5.3) and the unbiased Student estimator $A_{t,df=1}$ of formula 5.5, are compared with the desired the noncompliance probability $\alpha_{C>H}$=100-U%. If H/GM>1, U%>50 and

$$A_Z \leq 100\text{-}U\% \leq A_{t_{df\geq 1}}$$ Formula 7-1

are true, then a minimal sample size M_2=1,2,3,...? for the unbiased estimate of the compliance probability α exist.

If a value of df_2 exist, then the smallest sample size M_2 for an unbiased estimation of the compliance probability in a complete sample (M'=M) can be calculated using the formula 7.2 or 7.3 dependent if the sample is censored:

$$\text{if } \frac{M'}{M} = 1 \text{ then } M_2 = df_2 + 1$$ Formula 7-2

In a censored sample (with M' the number of results within the accuracy range):

$$\text{if } \frac{M'}{M} < 1 \text{ then } M_2 \text{ is the integer value for which :}$$

$$df_2 = INT\left(\frac{M_2 + INT\left(\frac{M'}{M} * M_2\right)}{2}\right) - 1$$

however if

$$df_2 > 2 * INT\left(\frac{M'}{M} * M_2\right) \text{ then } df_2 = 2 * INT\left(\frac{M'}{M} * M_2\right)$$ Formula 7-3

Because df_2, M' and M are known, M_2 can be derived iteratively. For GM/H<1 this routine is also useful to calculate the sample size for an unbiased estimate of $A_{C>H}$≥95% for being in non-compliance on the average.

The following situations are :

- 7.1 Compliance for every sample size.
- 7.2 Compliance for a specified number of samples.
- 7.3 Compliance for only a large number of samples.
- 7.4 Non-compliance for every sample size.

7.2 Two through 100 degrees of freedom

Tab 34c displays the increase of M_2 and the decrease $A_{t,df=i}$. It stops at the first integer value of M_2 for which $A_{t,df=i}<\alpha$ or if dfD99.

Tab 34b displays one of three messages:
- if $M_2<M$, then: "even for a smaller sample size than the current."
- if $M_2<M$, then: "for the current sample size."
- if $M_2>M$, then: "if sample size increases."

The minimal degrees of freedom df_2 is calculated using:

$$\text{if } t_{100\text{-}U\%,df=i+1} \leq \frac{LOG\left(\frac{H}{GM}\right)}{LOG(GSD)} * \sqrt{\frac{i+1}{i+2}} \leq t_{100\text{-}U\%,df=i}$$

$$\text{then } df_2 = i$$

formula 7-5

Contents - Index

7.3 More than 100 degrees of freedom

In practise it seems of little value to perform more than 100 measurements to show noncompliance with a limit value. The money involved which such a program could better be used to improve the working conditions. Screen 34c displays a situation that $A_{t,df}<\alpha$ for $99<df<\infty$.

The minimal degrees of freedom df_2 is calculated using:

$$\text{if } Z_{100-U\%} < \frac{\text{LOG}\left(\frac{H}{GM}\right)}{\text{LOG(GSD)}} * \sqrt{\frac{i+1}{i+2}} < t_{\alpha,df=100}$$

$$\text{then } df_2 = i > 100$$

formula 7-6

For GSD=1.20 and GM/H>0.7344, more than M2=50 samples are necessary to establish the desired probability α unbiased on the average. The bottom line of table 1 in Scheffers (1987) provides, for different GSD's, ratio GM/H values for which sampling is not useful. for $99<df<\infty$.

Contents - Index

7.4 Noncompliance

Tab 34d shows the state of being in permanent noncompliance: $A_{t,df} > \alpha$ for $df \rightarrow \infty$.

$$\text{if } \frac{LOG\left(\frac{H}{GM}\right)}{LOG(GSD)} < Z_{\alpha} \text{ then } df_2 \text{ does not exist}$$

formula 7-7

This situation is in noncompliance, even if the sample size is infinite. You can estimate that for α=5%, EXP(σ)=3.00 and EXP(μ)/H>0.1642 there is no sample size M2 that estimates α unbiased. The bottom line of table 1 of Scheffers (1987) provides for different GSD's, the EXP(μ)/H ratios over which measuring is ineffective.

Contents - Index

C How to get the program (running)

Follow the next steps to get a copy of HYGINIST running on your computer:

- C.1 Download HYGINZIP.EXE from the internet,
- C.2 Extract the HYGINIST setup files from HYGINZIP.EXE ,
- C.3 Remove earlier versions of HYGINIST for Windows from your computer,
- C.4 Install the program in your computer,
- C.4 Start the program .

Other items included here are:

- Register the program for continuous use,
- Stop execution.

Contents - Index

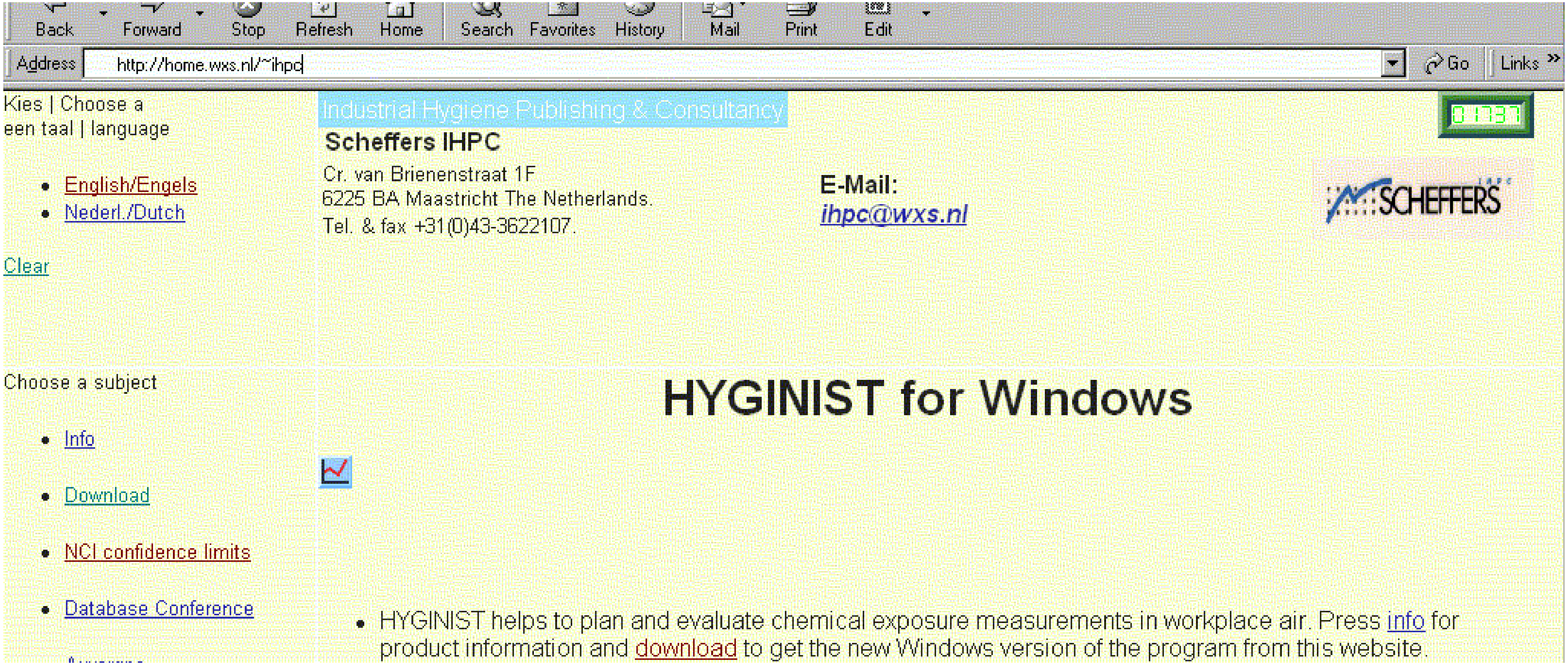
C.1 Download from the HYGINIST homepage

- Go to the HYGINIST internet homepage http://www.planet.nl/~ihpc
- Click in the frame on the left on download and subsequently in the frame on the right on HYGINZIP.EXE
- Now save HYGINZIP.EXE in a temporary directory and
- Wait while file is downloaded in the directory for temporary files on your system.

To do this in real press on http://www.planet.nl/~ihpc.

Continue with extracting the HYGINIST setup files from HYGINZIP.EXE

The HYGINIST internet site



Contents - Index

Download frame on the HYGINIST internet site

Choose a subject

- Info
- Download
- NCI confidence limits
- Database Conference
- Auvergne
- Auvergne weather

Download HYGINIST for Windows

HYGINIST for Windows is now distributed as shareware. Shareware means that :

- you can use the program during a 60 days free trail period,
- if you want to continu the use after that period, you need to register.

This 32 bits Visual Basic 5.0 version runs only in a Windows 95/98/ME/2000 or NT environment. HYGINIST for Windows contains the same functionallity as HYGINIST 2.2 (the DOS version) except for the comparison of two samples. It has many new features like:

- Windows object oriented,
- Paralel and interactive analysis,
- Shapiro and Wilk goodness-of-fit testing for censored samples,
- Exact confidence limits for and compliance probability with the Lognormal mean.
- Frequecy distribution plots.
- An extension on the non-Central Student Upper tolerance limits

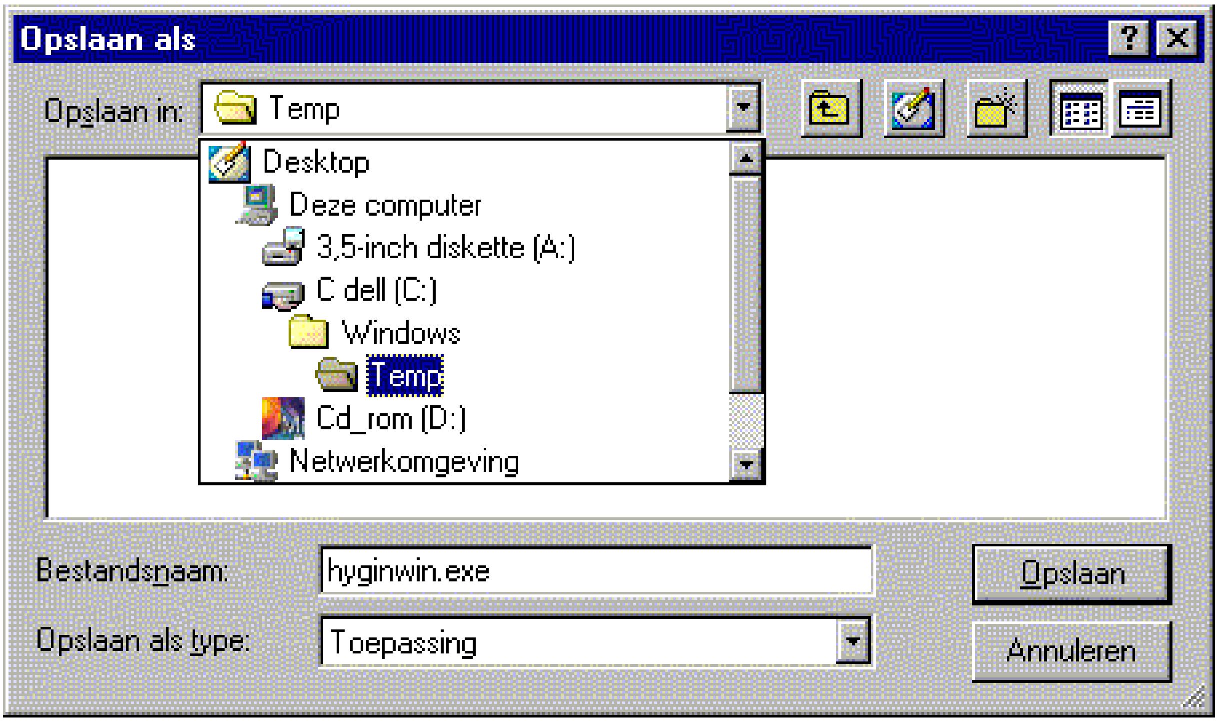
Click on HYGINZIP.EXE to start the download. The file with the current version (4.0.6#, date December 31, 2000) is 3.4 MB, so downloading may take some time. Start the self extracting WinZip program HYGINZIP.EXE. Continue by executing SETUP.EXE and follow the instructions. The distribution of HYGINIST 2.2 (the DOS version) has stopped. You can however still download HYGINIST 2.2 free of charge.

Contents - Index

Handling HYGINZIP.EXE

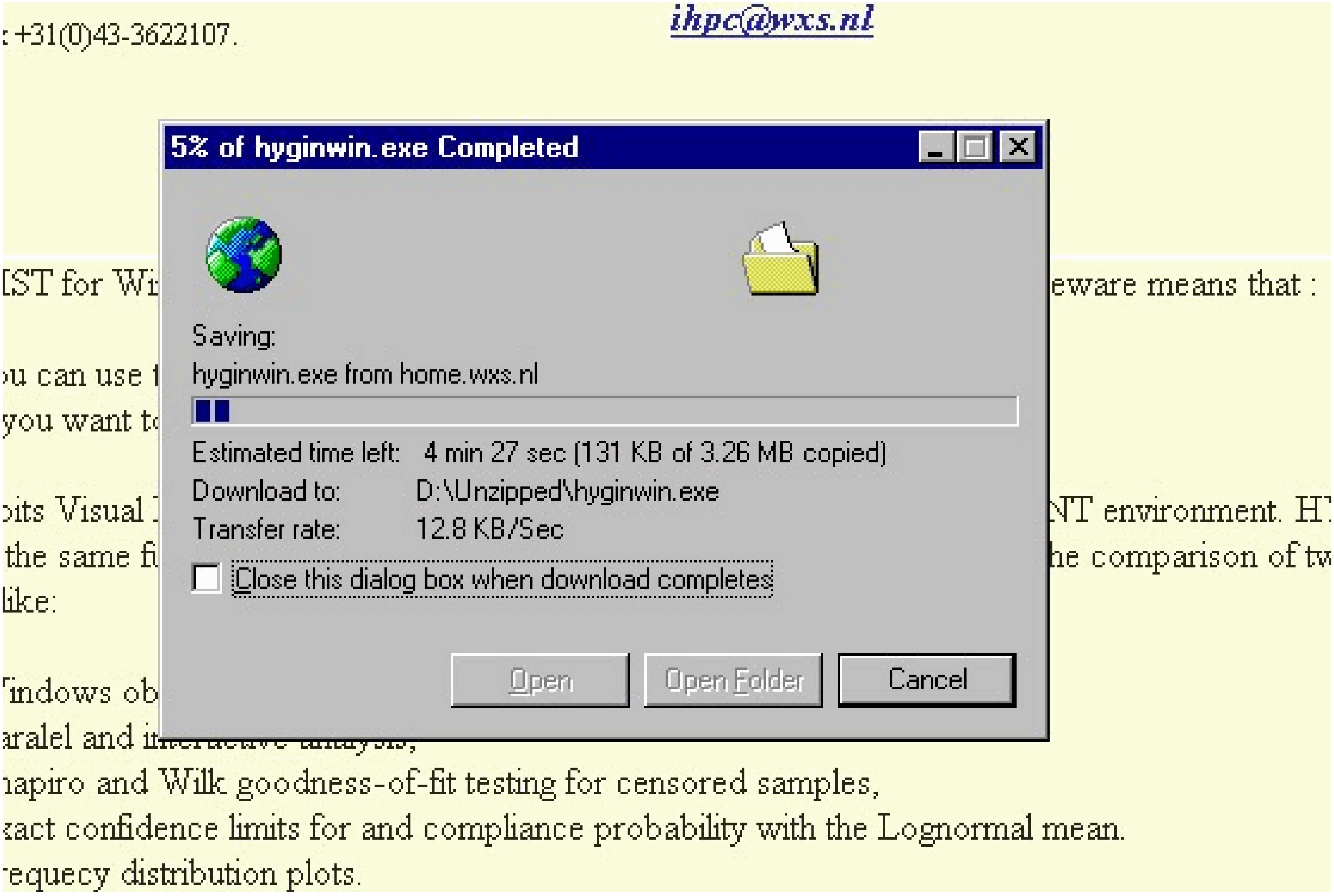
Contents - Index

Saving HYGINZIP.EXE



Contents - Index

Downloading HYGINZIP.EXE in your computer



C.2 Extracting HYGINZIP.EXE

Start the self-extracting program HYGINZIP.EXE.
It does the following:

- it will come with a suggestion to place the extracted files in a directory for temporary files
- Accept the suggestion or choose an alternative
- Keep the directory of choice name in mind.
- It will give you the number of files 

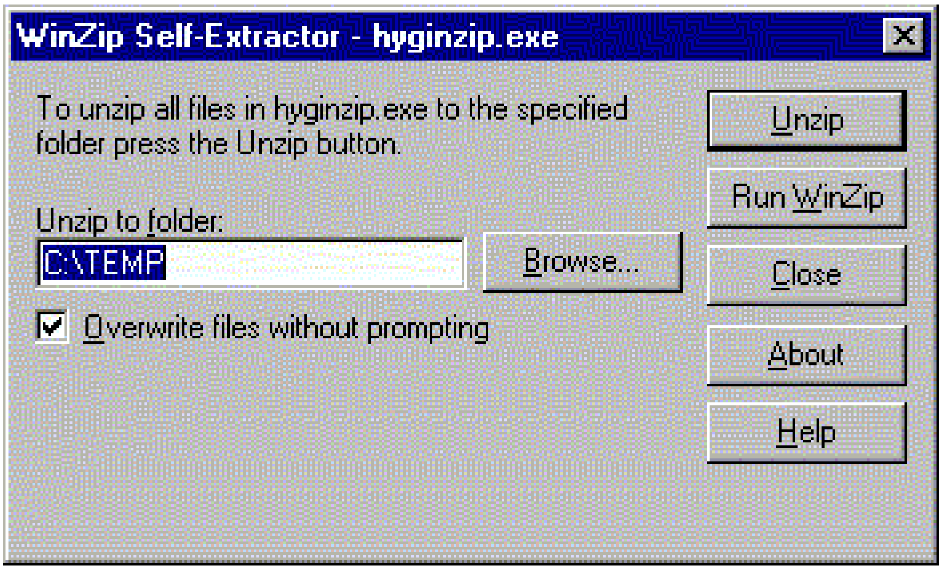
extracted

Continue with installing HYGINIST for Windows

If a HYGINIST for Windows (development) version was allready installed on this system remove it first.

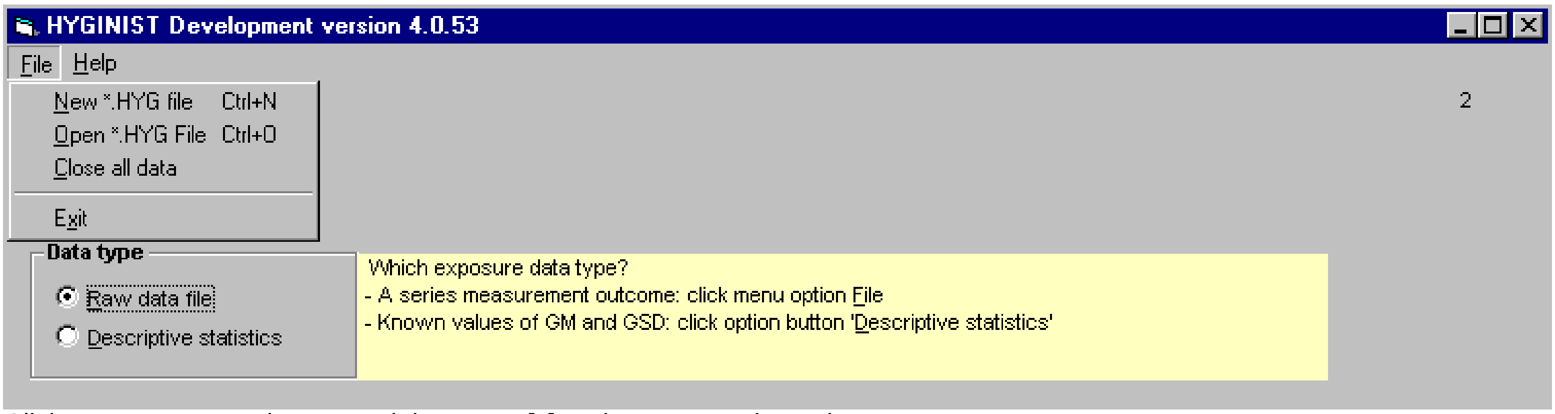
Contents - Index

Unzip

- 

Contents - Index

C.6 Stop and Exit the program



Click your cursor on the upper right corner [x] or the menu option exit.

E PROGRAM OPERATION

This chapter provides detailed information on:

- operating the program,
- screen lay-out,
- report making,
- on line help.

Contents - Index

E.1 Operating the program

In the next table the most important key functions and commands are explained.

<key> Comment See further
<Pause> Screen fixation. Press any key to continue -
string<Enter> Type and enter text (file name, dimension; W 8 characters) E.1.4, E.1.4.1 & E.1.4.2
digits<Enter> Type and enter a number E.1.4, E.1.4.1 & E.1.4.2
(sample size, exposure data, limit value, desired percentage)
<Delete> Delete typed or imported characters from the left hand side of the cursor E.1.4
<Backspace> Delete typed or imported characters from the right hand side of the cursor E.1.4
<Prt Sc> Sends current screen to clipboard E.3
←,↑,→,↓ Stirs reverse video block 2.2.4.1

Contents - Index

Goodness-of-fit (Summary)

(2) Sample air concentrations from similar exposure groups, will in general fit the Lognormal model if:

- sufficient data lie within the accuracy range of the measurement method,
- sampling procedures prevent random or systematic errors (outliers),
- the measurement coefficient of variation CV_t is smaller than the sample coefficient of variation ω/β, and
- the grouping interval ∇C is small compared to the arithmetic standard deviation ω,

Goodness-of-fit with the Lognormal distribution is assessed by:

- (2A) visual inspection of the probability plot of the logarithmic transformed data within the accuracy range.
- (2B) the W-test for Normality (Shapiro 1965), the most effective, complete sample test for shape.

Contents - Index

Extrapolation (Summary)

To deal with the common industrial hygiene practice of extrapolation from small (M<20) sample size sampling plans to general results, unbiased statistical estimators are calculated.

- (3A) In complete samples unbiased estimators are calculated for the geometric mean EXP(μ), the geometric standard deviation EXP(σ), the arithmetic mean β and the arithmetic variance ω^2 (Finney 1941, Laurent 1963).
- (3B) If at least two results lie inside and one outside the range of accuracy, then the estimators GM_g and GSD_g are derived from the Lognormal probability plot and the least squares linear regression through the data points within the range of accuracy (Gupta 1952). The M-M' data outside the accuracy range, contribute for one half to the total degrees of freedom df=(M+M')\2-1, however, for at most twice the number of the uncensored data df<2M'.

(4A) From a series full reference period Time Weighted Average (TWA) measurements, the "long-term non-compliance probability" (NIOSH 1977, page 65) or the "acute health hazard" (Hawkins 1991 page 56) can be assessed using:
- the standard Normal method (Leidel 1977),
- the unbiased method (Wilks 1941), or
- the Noncentral Student method (Tuggle 1982).

The standard Normal method is most efficient, but biased on the average and requires large sample size, small GSD's and very good Lognormal conformity. Wilks (for routine programs) and Tuggle (for first exposure assessment) have a higher sensitivity of in detecting non-compliance.

(4B) If the series exposure data are from within a reference period (e.g. grab or partial period, non consecutive measurements), then TWA compliance probability and confidence upper limit can be estimated using:
- the unbiased minimum variance method (Land 1971),
- the approximate, consistent Student test (Jahr 1987) used in HYGINIST version 2.2.

(4C) From a series TWA measurements "The chronic health hazard" (Hawkins 1991 page 56) or the "Cumulative dose" (Seixas 1988) are assessed with the same methods as in (4B).
(4D) If Ceiling compliance is defined as no more than α% exceeding the limit using grab or partial period sampling, then the confidence δ can be calculated using the Noncentral Student method.

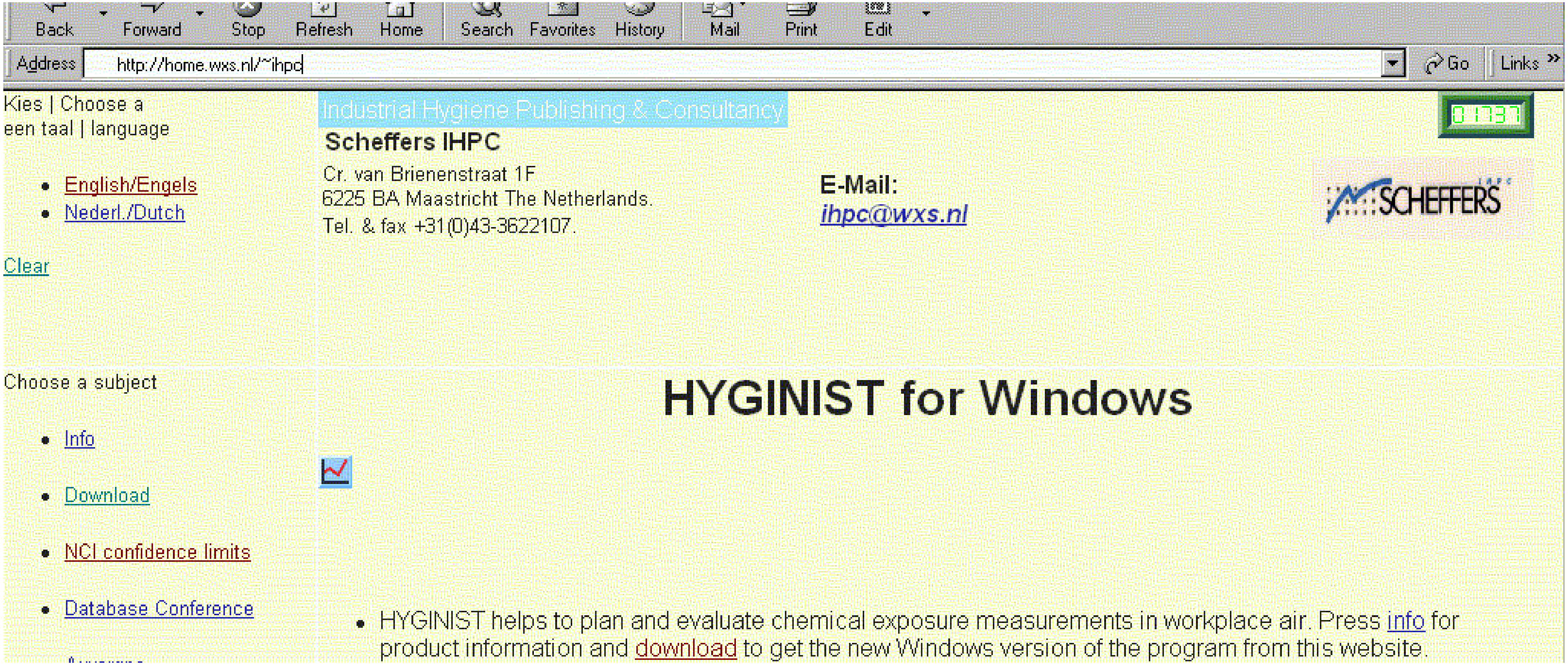
Contents - Index

HYGINIST email address

If you have any question please send an email to ihpc@planet.nl

HYGINIST Homepage

For additional information see the [HYGINIST homepage](#)



2.3 Using the data grid

2.3.1 Append

2.3.2 Edit

Contents - Index

2.3.3 Delete

Still no text

Contents - Index

3.1.2.1 Error messages

Tab "Plot" can display one of the following messages:

- "Less than 2 data points are within the limits.".
- "There is no variance between the limits": the relative difference of the highest and lowest value between the limits is below 10E-10.
- "There is too much variance between the limits": the ratio of the highest and lowest value between the limits is more than 10E+10.

Click "OK" and fill in a valid limit to continue.

Contents - Index

3.2.1 The W-test for uncensored samples

The picture box displays the chance A(W) that the (transformed) complete sample is drawn from a normal distribution.

Test statistic W is calculated using Formula B-1 . W values vary between 0 and 1. The probability distribution A(W) results in a value for goodness-of-fit between the 0 and 100%. Increasing values for A(W) suggest an increasing conformity with the shape of the Normal distribution.

In Example 16 and Example 17 W-values are tabulated and its influence the final Lognormal conformity conclusion is shown.

3.2.2 Censored sample

In case of a censored, logarithmic transformed sample Screen number 18b appears.

Screen 18b Goodness-of-fit for censored Lognormal sample

In the case of a censored non-Lognormal transformed sample Screen number 18c appears.

Contents - Index

4.2.1 Units of measurement

Not yet filled

4.2.1.1 Sampling units

Not yet filled

Contents - Index

4.2.1.2 Sampling duration

Not yet filled

Contents - Index

4.2.2 Exposure limit

Not yet filled

4.2.2.1 Value

Not yet filled

Contents - Index

4.2.2.2 Reference period

Not yet filled

4.2.3 Percentiles

Not yet filled

4.2.3.1 Confidence D

Not yet filled

4.2.3.2 Cover U

Not yet filled

Contents - Index

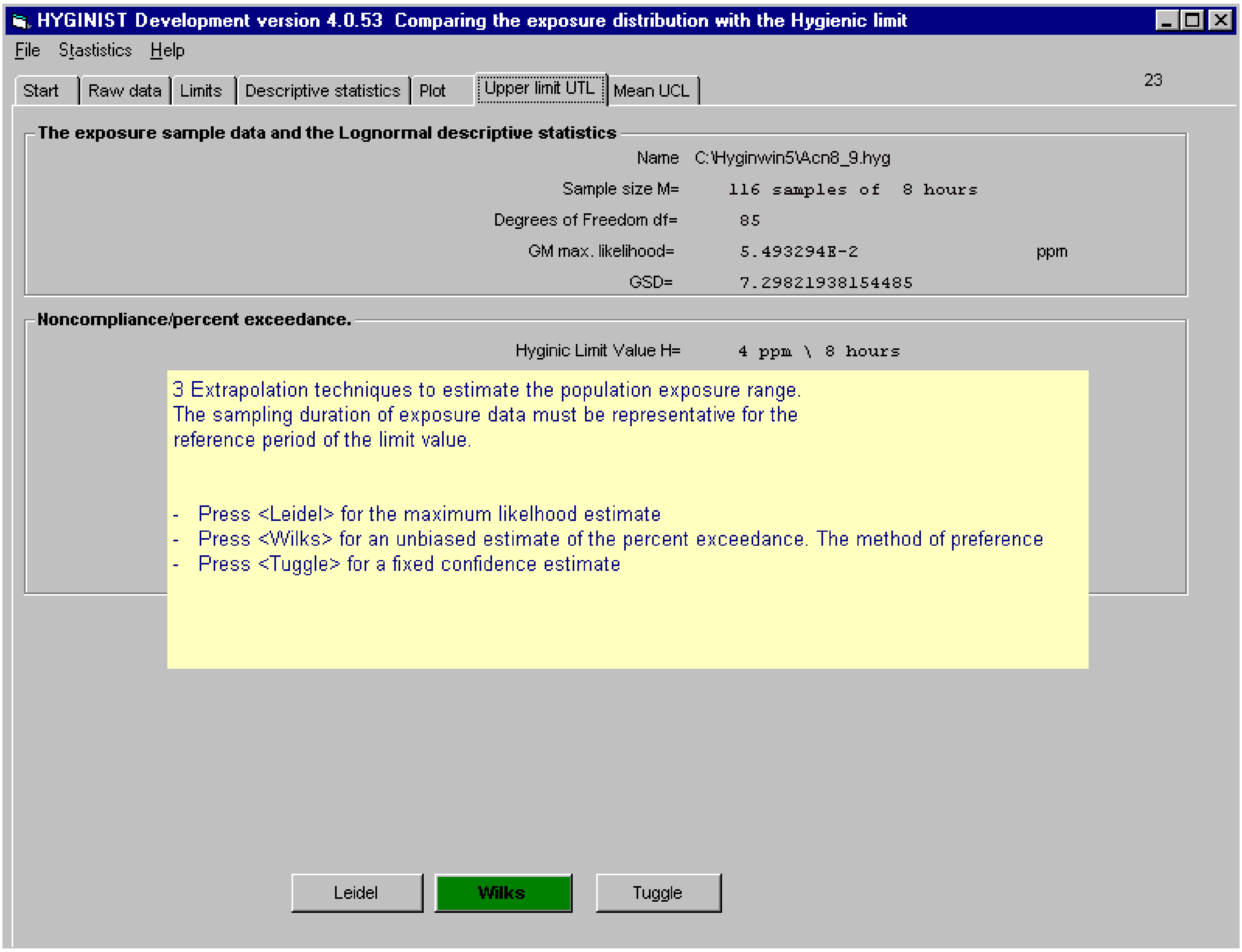
4.3 Lognormal frequency distribution

Not yet filled

Contents - Index

5.1.1 Comparing the 3 TWA extrapolation methods

To assess long-term compliance with hygiene limit H, 3 statistical extrapolation methods are presented and ranked (Screen 23a options 1, 2 & 3) in the increasing probability that H is considered as being an element of the exposure distribution. Use table 7 to choose the appropriate method.



The workplace air concentration distribution C in a homogeneous exposure situation is best described by the Lognormal distribution, If the population mean is μ and variance is σ^2, then the standard Normal deviate Z is the logarithm of C expressed in a standardized form:

Contents - Index

5.2.1.1. Nomograms

To test the sample based arithmetic mean against the hygiene limit H, plot the values of $_{10}$log(GSD) and:

- $_{10}$log(GM/H) in the nomograms of Bar-Shalom (1975 part II page 10 and 11), Bar-Shalom (1976 page 472) or Leidel (1977 page 58), for α=1% and α=5%, $_{10}$log(GSD)<0.5 (approximately GSD<3) and GM/H>.16,
- GM/H in the nomograms of Coenen (1978 page 405), for α=5%, 0.05<$_{10}$log(GSD)<0.95 (approximately 1.12<GSD<9) and GM/H>0.02 and M<1000,

and conclude:

- there is non-compliance with the limit value H,
- there is no decision possible; possible overexposure,
- there is compliance with the limit value H.

Contents - Index

6.2.3 Differences in location

Comparing sample GM's:

- is exclusively possible if sample GSD_2 and size M_2 are known,
- is performed with different test statistics for GSD<>GSD_2 and GSD=GSD_2.

The option button frame "GSDs from the same population" demands your conclusion on the population base of both GSD's. Are they from one or from two different populations?

- Compare GM's while GSD's are from same base (6.2.3.1)
- Compare GM's while GSD's differ significantly (6.2.3.2)

The power to detect differences in GM's is larger if the GSD's are considered as from the same base.

Contents - Index

B.1.3 Shapiro's W-test (censored sample)

For one-sided censored samples a adapted W-test is developed by Royston (1993,1995). The censoring must be:

- less the 80% and
- the sample size more than 20.

This test, extended to two sided censored samples, is included in HYGINIST for Windows. It was not available in HYGINIST version 2.2 and earlier.

The less effective{linkID=140140} Monte Carlo based test for Lognormality of Waters (1991) is easily expanded towards censored samples but this will lead to a mounting number of tables.

Contents - Index

B.5.1 Sample size using Wilks method

HYGINIST calculates the minimum sample size for long term compliance control and general exposure assessment based on the expected values for EXP(μ) and EXP(σ) and limit value H. Like Hawkins (1991 page 51) the robust Student distribution is used (see chapter 7). If formula 7-1 is true, then also:

$$Z_{100\text{-}U\%} \le \frac{\mathrm{LOG}\left(\frac{H}{GM}\right)}{\mathrm{LOG}(GSD)} * \sqrt{\frac{df_2+1}{df_2+2}} \le t_{100\text{-}U\%,\,df_2 \ge 1}$$

with $Z_{100\text{-}U\%} > 0$, $t_{100\text{-}U\%} > 0$ (because $U\% > 50\%$)

is true. Rearranging formula 5.24 so that df_2 becomes a function of the other parameters results in:

$$\frac{1}{\left(\frac{Z_{100\text{-}U\%}}{X}\right)^2 - 1} - 1 \le df_2 \le \frac{1}{\left(\frac{t_{100\text{-}U\%,df=1}}{X}\right)^2 - 1} - 1$$

$$\text{with } X = \frac{\mathrm{LOG}\left(\frac{H}{GM}\right)}{\mathrm{LOG}(GSD)}$$

Contents - Index

C.6.1 MenuExit

The HYGINIST program stops and asks you if you want to exit.

C.6.2 Exit Button

E.1.4 Entering characters or numbers

Text or numbers entered by keyboard, are completed if the object in which the are placed losses the focus.

- Comma's <,> in a series characters or digits are replaced automatically by a decimal point <.>
- Use <Backspace> or <Del> to remove undesired characters or digits.
- Text (name data file or dimension) may exceed 8 characters, however the text may be less visible. Upper and lower case characters are both accepted.

Contents - Index

E.1.4.2 Numbers

The decimal sign differs between and within the US and Europe. Comma's <,> in a series digits are replaced automatically by a decimal point <.>

Numbers <10000 and >= 0.1 are displayed with 9 digits at most, with a fixed decimal point.
####.######

Any number between 10000 and 1D+100 or between 1D-100 and 0.1 is displayed with at most 8 digits and with a fixed decimal point combined with an exponent `D±NN`. The real value of the number is the mantissa multiplied by 10 raised to the ±NN power.
##.######D±NN

Numbers >= D+100 or <=1D-100 who do not exceed the extremes, are displayed with at most 7 digits and with a fixed decimal point combined with an exponent `D±NNN`. The real value of the outcome is, the mantissa multiplied by 10 raised to the ± NNN power.
##.#####D±NNN

Contents - Index

Context sensitive help

Tool tips

Contents - Index

Internet

Visit the HYGINIST homepage

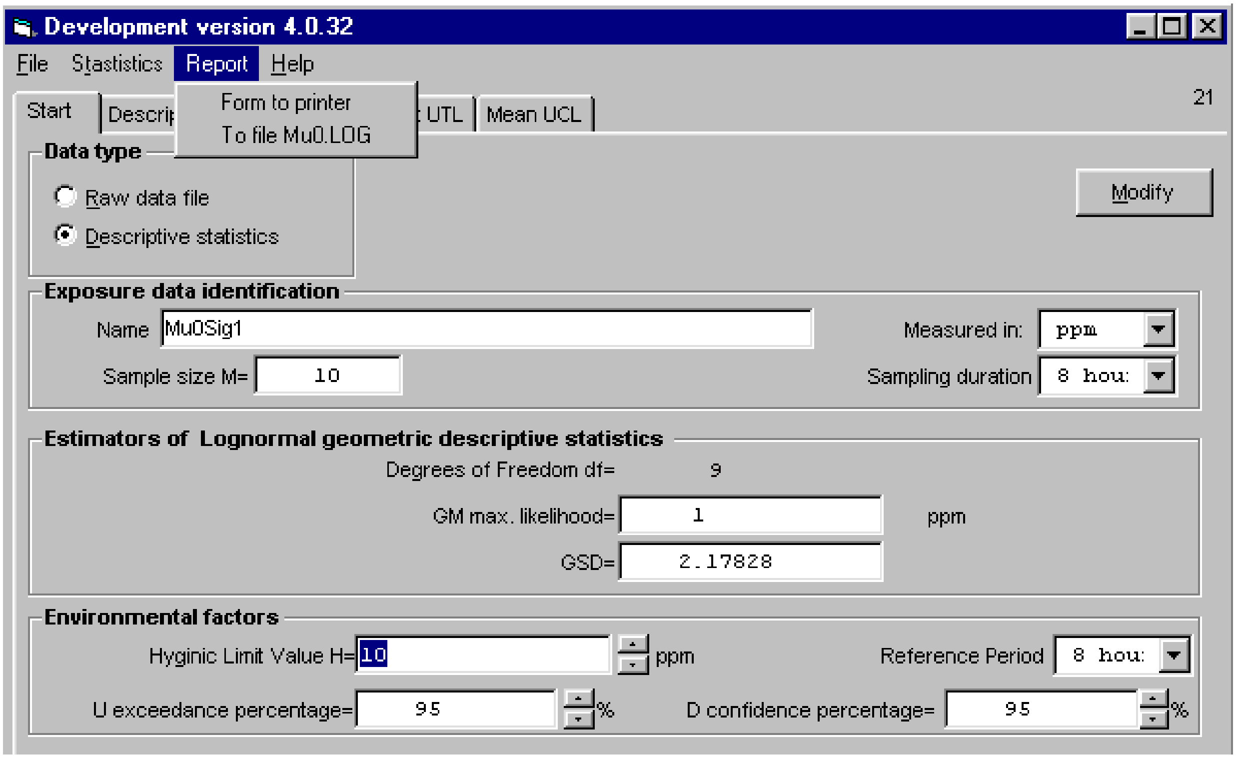
E.3 Report facility

All HYGINIST screens are easily incorporated in text editor programs. Since standard text and picture processing programs have extended facilities to manipulate text and figures in reports HYGINIST screens are only made available for these programs

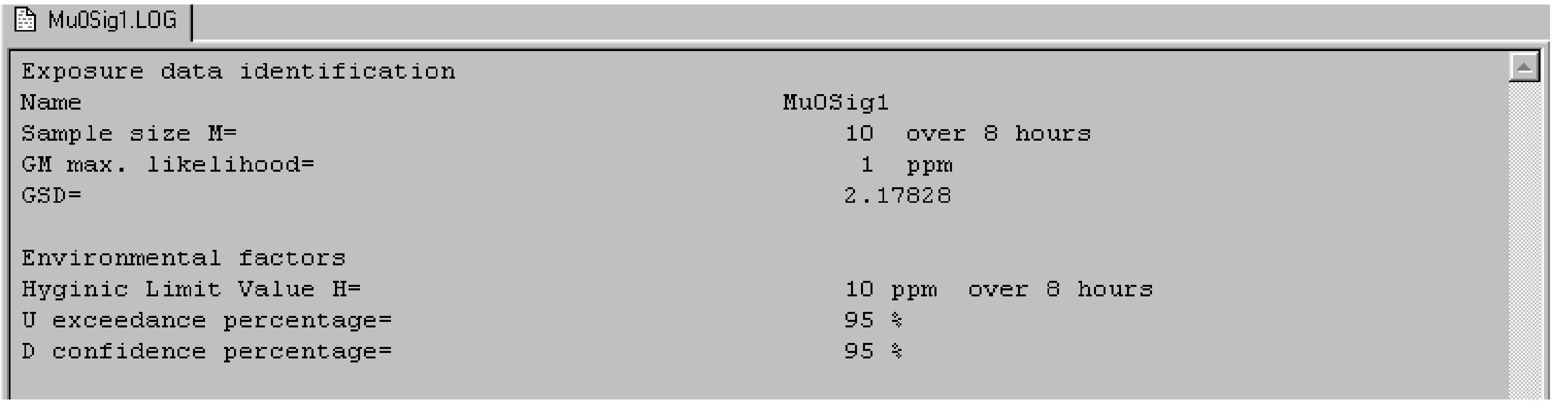
Contents - Index

E.3.1 *.LOG report text file

If Report is displayed in the menu bar, then you can append the information of the current screen to the TEXT data file *.LOG.



If HYGINIST.LOG does not exist in the current directory, then it is created automatically.
To report exact figures, use HYGINIST.LOG. The file can be imported in any word processing program.


MuOSig1.LOG

```
Exposure data identification
Name                                              MuOSig1
Sample size M=                                        10   over 8 hours
GM max. likelihood=                                    1   ppm
GSD=                                                  2.17828

Environmental factors
Hyginic Limit Value H=                                10 ppm  over 8 hours
U exceedance percentage=                              95 %
D confidence percentage=                              95 %
```

Contents - Index

E.3.3 Printer

Contents - Index

F References

F.1 Computer program and quantitative evaluation

F.2 The example data in Table 1

F.3 Further readings

Contents - Index

F.2 The example data in Table 1

Albrecht W.N., Burr G.A., Neumeister C.E. Sampling and analytical method for workplace monitoring of aspartame in air. Applied Industrial Hygiene 4 (1989) 217-221. Table II Weight-out.
Bar-Shalom Y., et al Handbook of Statistical Tests for Evaluating Exposure to air contaminants. HEW Publication No. (NIOSH) 75-147 Washington: U.S. Department of Health, Education and Welfare (1975) I-25.
Bilan R.A., Haflidson W.O., McVittie D.J. Assessment of isocyanate exposure during spray application of polyurethane foam Am. Ind. Hyg. Assoc. J. 50 (1989) 303-306. Table 1 sprayer.
Boleij J., Heederik D., Kromhout H. Karakterisering van blootstelling aan chemische stoffen in de werkomgeving. Pudoc Wageningen (1987) 62 & 84.
Booher L.E. Lead exposure in a ship overhaul facility during paint removal. Am. Ind. Hyg. Assoc. J. 49 (1988) 123.
Cohen A.C. Three-Parameter estimation. In: *STATISTICS: Textbooks and Monographs Volume 88 Lognormal Distributions. Edited by Edwin L. Crow and Kunio Shimizu.* New York: Marcel Dekker (1988) 113-137.
Conover W.J. Practical Nonparametric Statistics. Texas Tech University. John Wiley & Sons (1980) 195, 365.
Dewell P. Some applications of statistics in occupational hygiene. BOSH Technical Handbook Series No. 1 (1989) 24, 42 and 44.
Gupta A.K. Estimation of the mean and the standard deviation of a Normal population from a censored sample. Biometrika 39 (1952) 271.
Gustafson T.L. TRUE EPISTAT Manual chapter 6. Epistat services Texas (1989) 17.
Hald, A. Statistical Theory with engineering applications. par.. 6.9.: The truncated Normal Distribution. Wiley (1952) 151.
Hawkins N.C., Norwood S.K., Rock J.C. (ed.) A strategy for Occupational Exposure Assessment. Akron (Ohio): AIHA (1991) 104, 105 and 110.
Leidel N.A., Bush K.A., Lynch J.R. Occupational exposure sampling strategy manual. DHEW NIOSH-137 (1977) 56, 61, 63, 67, 103 and 104.
Owen W.J., DeRouen T.A. Estimation of the mean for Lognormal data containing zeroes and left-censored values, with applications to the measurement of worker exposure to air contaminants. Biometrics 36 (1980) 707-719.
Post W.K. Schatten van blootstellingen: Een onderzoek naar het nut van kwalitatieve blootstellingsschattingen en de (on)zin van indicatieve metingen. LUWageningen Rapport Gezondheidsleer 382 (1989) bijlage 11 vervolg, Styreen, funktie inlegger.
Potts D.L., Craft B.F. Occupational exposure of Veterinarians to waste anaesthetic gases. Applied Industrial Hygiene 3 (1988) 134.
Sarhan A.E., Greenberg B.G. Contributions to Order Statistics. New York: John Wiley (1962) 212.

Scheffers T.M.L., Jongeneelen F.J., Bragt P.C. Actual health-related hazard from solvents to maintenance painters. Ann. Occup. Hyg. 29 (1985) 198.
Schneider H. Truncated and censored samples from Normal populations (Statistics, textbooks and monographs ; vol 70) New York: Marcel Dekker (1986) 69 and 88 {a $_{10}$Log transformation of the data of Gupta (1952)}, 70 and 224.
Swaen G.M.H., Bloemen L.J.N., Twisk J., Scheffers T., Slangen J.J.M., Sturmans F. Mortality of workers exposed to acrylonitril JOM 34 (1992) 801-809.

Contents - Index

F.3 Further readings

AIHA Quality assurance manual for industrial hygiene chemistry. American Industrial Hygiene Association. Prepared by the Analytical Chemistry Committee of the AIHA. (1988).
Aitchison J., Brown J.A.C. The Lognormal distribution. Cambridge University Press (1957) Edition 1976.
Ali M.M., Chan L.K. On Gupta's estimates of the parameter of the Normal distribution. Biometrika 51 (1964) 498-501.
Cohen A.C. Simplified estimators for the Normal distribution when samples are single censored or truncated. Technometrics 1 (1959) 217-237.
Corn M., Esmen N.A. Workplace exposure zones for classification of employee exposures to physical and chemical agents. Am. Ind. Hyg. Assoc. J. 40 (1979) 47-57.
Hall I.J., Sampson C.B. One-sided tolerance limits for a Normal population based on censored samples. J. Statist. Comput. Simul. 2 (1973) p 317-324.
Hering S.V. Air sampling instruments for evaluation of atmospheric contaminants. 7th edition. Am Conf. Governm. Ind. Hygienists Cincinnati (1989).
Hertig B.A. Work in hot environments: Threshold Limit Values and proposed standards. In: Industrial Environmental Health, The Worker and the Community. New York: Academic Press (1975) 219-231.
Kennedy W.J. Gentle J.E. Statistical computing. Chapter 5: Approximating probabilities and percentage points in selected probability distributions. Statistics: textbooks and monographs. Vol 33 (1981) 61-132.
Kromhout H., et al Agreement between qualitative exposure estimate and quantitative exposure measurements. Am. J. Ind. Medicine 12 (1987) 551-562.
Lewis H., et al Aerodynamics of the human microenvironment. Lancet 1 (1969) 1273-1277.
Myers R.A., Pepin P. The robustness of Lognormal-based estimators of abundance. Biometrics 46 (1990) 1185-1192.
Pirsig R.M. Zen and the art of motorcycle maintenance. An inquiry into values (1974).
SKC Guide to air sampling standards, Blandford Forum, UK, (1991).
Selvin S. A graphical estimate of the population mean from censored Normal data. Applied Statistics 25 (1976) 8-11.
Theune J.A. Comparison of power for the d'Agostino and the Wilk-Shapiro test of Normality for small and moderate samples. Statistica Neerlandica 27 (1973) 163-169.
Thöne H., A table for estimating the mean of a Lognormal distribution. Am. Stat. Assoc. J. 64 (1969) 632-636.
Zielhuis R.L., Wibowo A.A.E. Standard setting in occupational health: Philosophical issues. Am. J. Ind. Med. 16 (1989) 569-598.

Contents - Index

G Symbols and abbreviations

Greek letters α,β,γ,δ... are population parameters.
Latin letters A,D.... are sample estimates.

Symbol name description
α alpha Change, probability,
A Sample estimator of chance ?
A_{gem} Arithmetic mean of a series probabilities
A(W) Probability distribution of Shapiro's W-test (1965)
AM Finney's estimator of β
AM_{am} Direct, simple or arithmetic estimator of β
AM_{ml} Maximum likelihood estimator of β
β beta Arithmetic mean of the Lognormal distribution. First moment
BLU Best linear unbiased estimators (Schneider 1986 p57)
c_i Measured concentration in workplace air in period i
C_i True concentration in the workplace air in period i
C_T True concentration in workplace air in reference period T
C(δ,α) Calculated concentration overlying at least U% of the Lognormal distribution
Ç(δ,α) Estimator of C(δ,α)
CI Confidence Interval (two sided upper and lower limit)
CV_t Sampling an analytical measurement coefficient of variation (Leidel 1997 p78)
δ delta Confidence, reliability
ΔC Grouping interval
D Sample estimator of the confidence δ
d Noncentrality parameter
df Degrees of freedom
E() epsilon The expectation, the theoretical arithmetic mean of the probability distribution
EXP() De exponential distribution with base e=2.71828...
EXP(μ) Geometric mean of the Lognormal distribution
EXP(σ) Geometric standard deviation of the Lognormal distribution
F Deviate of the Fisher or variance ratio distribution
GM Direct sample estimator of EXP(μ)
GM_U Unbiased sample estimator of EXP(μ)
GM_G Regression sample estimator of EXP(μ)
GSD Unbiased sample estimator of EXP(σ)
GSD_G Regression sample estimator of EXP(σ)
H Industrial hygiene limit value
SEG Similar exposure group
INT() Integer, a function to establish the value before the decimal point
$\Lambda(\mu,\sigma^2)$ Lambda Lognormal distribution with descriptive statistics EXP(μ) and EXP(σ)
LCL Lower confidence limit of a descriptive statistic
UTL Lower tolerance limit of a distribution
LOG() The natural logarithm with base e=2.71828...
μ mu The mean of a Normal distribution
M The total sample size (censored and uncensored data)

M' The number of results within the detection limits
M_τ The population size (confined or expanded)
ML Maximum likelihood
MM Method of moments (Schneider 1986 p39)
MML Modified maximum likelihood (Schneider 1986 p97)
N(0,1) Standard Normal distribution with μ=0 and σ=1
OTL One-sided tolerance limit
Π Product of a series results
R_i Rankit, normal order statistic, standard normal order expectations
RML Restricted maximum likelihood (Schneider 1986 p100)
RSD see CV_t and ω/AM
σ sigma Standard deviation of the normal distribution
s Sample estimator of σ
S_i Coefficient for the normal W-test
Σ Sum of a series results
SQR() The square root
t Deviate of the Student distribution
t_i A sampled period within reference period T
t_j An unsampled period within reference period T
T The duration of the reference period
$T\sigma_i$ The sum of all sampled periods within the reference period T
T_d Deviate of the Noncentral Student distribution
? tau The sum of all reference periods T
TI Tolerance Interval (two sided upper and lower limit)
TWA The arithmetic, time weighted average concentration within the reference period
U% The desired percentage for confidence and tolerance limits
UCL Upper confidence limit of a descriptive statistic
UTL Upper tolerance limit of a distribution
W The probability density function of the analysis of variance test for Normality (Shapiro 1965)
ω_2 Variance of the Lognormal distribution. Second moment
ω/AM The relative arithmetic standard deviation
x_i log(c_i) The natural logarithm of concentration c_i
Direct estimator of the arithmetic mean μ of the Normal distribution
χ Chi Square root of the deviate of the Chi-square distribution
Z Deviate of the standard Normal distribution *N*(0,1).

Contents - Index

Detection range

see range of accuracy

Contents - Index

Range of accuracy

The values for which the measurement values are considered to be reliable.

Contents - Index

HEG

- A group of workers with identical probabilities of exposure to a single environmental agent (Hawkins 1991 page 5).
- A group of employees who experience agent exposures similar enough that monitoring of any worker in the group provides data useful for predicting exposures of the remaining workers (Hawkins 1991 page 160).

see also Similar Exposure Group

Contents - Index

Population

Collection of elements on which the conclusions of the statistical evaluation are related to

Contents - Index

Random drawing

- Taking an element from a population with a method that is independent of all properties of the element

Contents - Index

Reliability

See confidence

Robust

The property of a test that it works well for a wide variety of population types (Snedecor 1980 p 135)

Contents - Index

Rankit

The expectation of the standard normal ordering (Fisher 1938)

Contents - Index

Reference period

The specified period of time stated for the limit value of a specific agent (prEN 689)

Contents - Index

Sample

A series drawings from a population

Contents - Index

Scale

Variation

Contents - Index

SEG

See Similar exposure group

Contents - Index

Selectivity

Degree of independence from interferents (prEN 482)

Skewed distribution

An asymmetrical distribution (Kendall)

Sustainment

- he capacity of individuals or groups to bear exposure

Contents - Index

Truncated distribution

A distribution with an unknown proportion result outside the measurement reach (Hald 1949)

Contents - Index

Upper tolerance limit

- The upper boundary containing at least the desired fraction of the population

Contents - Index

Unbiased estimator

An estimator with the characteristic that the expectation equals the parameter for every sample size.

Contents - Index

Workplace

The defined area or areas in which the work activities are carried out.

Contents - Index

J About the author

HYGINIST has been quite a solo project. Research, developing and writing the approach, the program and the enplaning manual started in 1985. I was born in 1953 in Voorburg (the Netherlands). Awarded in Environmental Health and Hygiene in 1979 at the Agricultural University in Wageningen based on a study on the concentrations of tetrachloroethylene in the breathing air of neighbors of dry-cleaning shops. In 1979 I joined the occupational health research organization TNO-MBL. There I performed a two-year health hazard survey among painters working in the Dutch construction industry. In 1981 I joined the DSM Chemical company to co-ordinate the industrial hygiene activities on a site of 750 ha with 51 chemical plants and 7000 employees. I followed several post-doc courses on biostatistics and epidemiology and performed and co-operated in several epidemiology studies. The exposure assessment in health and hygiene surveillance became my main branch of interest and from both that and the passion for industrial hygiene, statistics, computers and communication, the computer program HYGINIST evolved. In 1989, I joined the Dutch Occupational Hygiene Society's Limits and Measurement Methods Group, of which I was the chair between 1991 and 1997. I am full member of the American Industrial Hygiene Association AIHA's Exposure Assessment Strategies Committee and the WEEL committee. In 1994 I incorporated the industrial hygiene education, consultancy and publishing activities in Scheffers IHPC I'm living with wife and daughter in Maastricht, one of Europe's finest cities.

Contents - Index

Table 6 Characteristics of the complete sample estimators of the Lognormal descriptive statistics

Estimator Description and properties additional to Table 5

M' the number of results between

Formula 4-7 upper and lower limit $M' = ul - ll$

df the number of degrees of freedom

$$\text{If } df \le 2 \times M' \text{ then}$$

$$df = \text{INT}\left(\frac{M + M'}{2}\right) - 1$$

$$\text{else if } df > 2 \times M'$$

$$\text{then } df = 2 \times M'$$

Formula 4-8 of a censored sample: M'<M

GSD_g Geometric Standard Deviation estimated
Formula 4-9 using the data pairs between the accuracy limits.

$$GSD^g = EXP\left(\frac{\sum_{j=ll}^{j=ul} R_j * x_j - \frac{\sum_{j=ll}^{j=ul} R_j \sum_{j=ll}^{j=ul} x_j}{M'}}{\sum_{j=ll}^{j=ul} R_j^2 - \frac{\left(\sum_{j=ll}^{j=ul} R_j\right)^2}{M'}}\right)$$

GM_g Geometric mean estimated using the
Formula 4-10 data pairs between the accuracy limits.

Table 6 Characteristics of the complete sample estimators of the Lognormal descriptive statistics

$$GM^{g} = EXP\left(\frac{\sum_{j=ll}^{j=ul} \left[x_j - LOG(GSD^{g}) * R_j \right]}{M^{r}} \right)$$

Contents - Index

Table 9 Statistical properties of 3 methods which assess the effectiveness of control

Method-> Standard Normal   Unbiased Noncentral Student

Characteristic Chapter 5.1.2 Chapter 5.1.3 Chapter 5.1.4

Expectation of A%: E(A%)>α E(A%)=α E(A%)<α
- increase of the bias if: GSD ↑, Unbiased GSD ↓,
H/GM ↑ or df ↓ H/GM ↓, or df ↑

Efficiency of Variance High Moderate Lowe
σ(A%) estimate

Shape of A% Skewed, but asymptotic Normal if GSDº1 or H/GMº1
 - outlier incidence Frequent and extreme Moderate lowest
 - kurtosis shape Tapering Moderate flattest
Leptokurtosis platykurtosis

Confidence δ of finding
non-compliance.
 If Ca=H, then $0<\delta_{A>=\alpha}<50\%$ $50<\delta_{A>=\alpha}<100\%$ $d_{A>=\alpha}=U\%$
 if df->0, then $\delta_{A>=\alpha}$->0% $\delta_{A>=a}$->100% $\delta_{A>=\alpha}=U\%$

Contents - Index

Table 11 Methods to test Ceiling limits using TWAshort period

Lognormal conformity kC>H Sample size M Statistical method Chapter
Rejected D0 and Wp*M >1/p (cumulative) Binomial/Poisson 5.3.1
Dacceptable n.r W1/p adapted Binomial/Poisson 5.3.2
Dacceptable n.r >2 Lognormal Noncentral Student 5.3.3

Contents - Index

Table 12 Notations and terms used for inference test statistics.

Name Population sample chapter
parameter estimate
Two-sided probability, GSD = GSD2I αGSD=GSD2 AGSD=GSD2 6.2
nference of two sample Geometric Standard Deviations
Two sample difference in scale
Two sample analysis of variance

Two-sided probability, GM = GM2 αGM=GM2 AGM=GM2 6.2.3
Inference of two sample Geometric Means
Two sample difference in location

Two-sided probability, GSD = EXP(σ
)Inference with population Geometric Standard Deviation αGSD=EXP(σ) AGSD=EXP(σ) 6.3.1

Two-sided probability, GM = EXP(μ) αGM=EXP(μ) AGM=EXP(μ) 6.3.2
Inference with population Geometric Mean

Table 13 Effective control of occupational exposure risks

```
+----------+   +------------+
|         1|   | Limit     2|
| hygienic +-->| Value      +--+  +--------------+
| hazards  |   +------------+  +->|  Exposure   4|yes/unknown
+--|-------+   +------------+  |  |  compliance  +---+
   |   |       | Exposure  3+--+  +------|-------+   |
   |   |       +-----|------+            |           |   +-------------+     +------------+
   |   |             |         +---------+no         |   |            8| yes | Effective 9|
   |   |       +-----+------+  |  +--------------+   +-->| Sustainment +--->| Control     |
   |   |       |           5|<-+  |mental &     7+---|   |             |     |            |
   |   |       | Control    +---->|physical      |   |   +-------------+     +------------+
   |   +------>| measures   |     |resistance    |   |           |no               no |
   |           |            |     +--------------+   |           |                    |
   |           |            |     +--------------+   |           |                    |
   |           |            +---->|Co-exposure  6+---+           |                    |
   |           +-----|------+     +--------------+               |                    |
   |                 |                                           |                    |
   +-----------------+----------------------<--------------------+                    |
   +-----------------------------------------<----------------------------------------+
```

Contents - Index

Example 20 Correcting exposure variability for a fixed background

*.HYG file Description

LEIDL103 The Lognormal probability plot on M=12 successive Hydrogen fluoride concentrations (Leidel 1977 p103-104) clearly shows the fixed background exposure level of about 0.1 PPM additional to Lognormal variation.

LD103_10 If 0.1 PPM is subtracted from all results of LEIDL103.HYG, then GM=.18 PPM and GSD=9.75: exactly the same as the values calculated by Leidel (1977 p103). The probability plot, however suggest the need to introduce a lower limit of .01 PPM and this results in rankit estimators of M'=10, GM_g=0.16 PPM and GSD_g=13.9: somewhat more accurate Than the estimators read by Leidel (1977, figure I-4 page 104) from the Probability plot GM=.16 and GSD=(84% value)/(50%value)=12.8.

Contents - Index

Example 22 How to establish sample size from GM and GSD

6 TWA8 hour total dust samples from a hypothetical homogeneous population (with known μ=-0.1, σ=1) results in the GM and GSD values which are displayed in Screen 7ab. The gravimetric total dust TWA8 hour method has a CVt=0.25 (Leidel 1977 table D-1) which is smaller than the displayed sample relative standard deviation w/AM=1. Thus, Lognormality is not overshadowed by measurement error. The minimal sample size for long-term compliance control with H=10 mg/m3 and U=95% is estimated in chapter 7.

Contents - Index

Example 23 Exceeding the limit (censored non-parametric)

*.HYG file Description

HALD151 If one wishes to know what percentage of diameters of rivet heads (Hald 1952) are below 13.3 mm, then Screen 21a will respond "Between 0 and 38%" because of the lower detection limit of 13.40 mm

SCHNE224 The upper accuracy limit at 6 Time Units (Schneider 1988 page 224) for the items under stress, inhibits the Nonparametric estimation of the number of items exceeding 10 TU (see Figure 5).

Example 24 Using 95% as default

Nearly every application of statistics uses 95% as default extrapolation value, or its complement 5% as significance level:

- Leidel (1977 page 69) uses in the compliance control $\alpha>0.05$ to indicate that control measures are necessary.
- The European-based Dutch legislation on Vinylchloride requires 95% confidence that the average annual concentration in the workplace air, calculated from sequential measurements, is below the limit value TWA1 year.
- According to Leidel (1977 page 118) an employer should try to attain 95% confidence that no more than 5% of the employee days are over the limit value.
- The one-sided tolerance (OTL) approach of Tuggle (1981 page 497) questions "Are we 95% confident that less than 5% of all exposures received in the workplace are above the PEL?"
- In the 8 hour labour service, $\alpha=0.05$ corresponds with a 24 minutes period and with 1 out of 20 shifts.

Example 25 Two-sided interval

In order to calculate the two-sided 95%:

- confidence interval of the arithmetic mean by using the method of Jahr while running the procedure "Classification for TWA limit/Cumulative dose." (Screen 24, par. B.2.),
- tolerance interval running the procedures "effectiveness of control measures" from NIOSH or Wilks (Screen 25 and Screen 26),

the desired percentage should alternate between U=2.5% and U=97.5%

Contents - Index

Example 34 Estimation of the average daily dose

*.HYG file Description

ACN8_9 In a retrospective cohort mortality study (Swaen 1992) among Dutch workers occupational exposed to Acrylonitrile (ACN), the average daily dose was assessed using the exposure control chart (Figure 17) of the similar exposure group (=SEG) ACN production workers. The M=116 PAS $TWA_{8\ hour}$ ACN, sampled in 1977 and 1978, fit the Lognormal model (see Figure 18) when a lower detection limit of LL=0.05 PPM is introduced. The M`=56 uncensored data provides df=85 degrees of freedom. Using the rankit estimators GM_g=0.05 PPM and GSD_g=7.3, the average daily dose in 1977/1978 is, for this SEG, estimated as $AM_{1977\text{-}1978}$=0.37 PPM (90% confidence range of 0.25-0.53 PPM).

VINCHL91 The M=37 results of the Vinylchloride PAS sampling in a SEG PVC polymerisation shift workers (see Example 29) sampled in 1991 fit the Lognormal model (see Figure 16). With df=36, GM=0.36, GSD=2.90 the average daily dose is estimated as AM_{1991}=0.60 PPM.
The compliance against the European Community exposure limit of HEC=3 PPM (a one year average of all $TWA_{8\ hour}$'s) is $A_{(\beta>H)}$=10-9, indi cating that the long term control measures are effective for this SEG.

HAW117 From 14 random TWA8 hour PAS total dust above LL=1.4 mg/m3 the chronic health hazard was assessed (see Example 57).
With AM=2.2 mg/m3 and H=10 mg/m3 the approximate probability that β exceeds limit H is $A_{(\beta>H)}$=7*10-12%. The two-sided 95% confidence interval including β is 2.0-2.47 mg/m3.
Though the algorithms for the interval calculation by Hawkins (1991 page 130-133 approach 1 and 2) are quit different, the result (Hawkins 1991 page 134 table III.2 two last lines) are comparable (Normal approach 1.97-2.43 mg/m3, Lognormal approach 1.93-3.42 mg/m3).

- From formula 5.13, the Student table (Ciba-Geigy 1980) and GSD=2.71828 it follows that:
for df=1, $Ç_{\beta,\delta>=95\%}/\beta$=552.1
for df=9, $Ç_{\beta,\delta>=95\%}/\beta$=6.311.
If sample size increases, then $Ç_{\beta,\delta}/\beta$ decreases towards unity.

Contents - Index

Example 35 Estimation of grab sample based TWA

*.HYG file Description
OWEN716 From GMg=.88 PPM, GSDg=6.81 and df=10 of the series grab sample airborne Chlorine , the two-sided confidence interval is 1.3-11.7 PPM, indicating real overexposure against H8 hour=1 PPM (Owen 1980 p716).

BAR_SI25 From GM=47.5 PPM and GSD=1.675 on carbon monoxide grab samples a TWA8 hour=53.0 PPM is estimated. The probability that β=TWA8 hour exceeds H8 hour=50 PPM is AßDH=59.4% (90% confidence interval 33.3-84.3 PPM) indicating possible overexposure (Bar-Shalom 1975 page I-25).

LEIDEL61 From 35 direct reading Ozone concentrations, the estimators of the untransformed descriptive statistics μ and σ are =.0831 and s=.023 PPM. The one-sided 95% upper confidence limit is calculated as UCL=.089 and the situation was concluded to be in compliance with the H8 hour=.1 PPM (Leidel 1977 p61-62). Using formula 5.12 the 95% UCL of β is Çβ,U%=95=0.090, which leads to the same conclusion.

Contents - Index

Example 36 Compliance with a ceiling limit value (distribution free)

*.HYG file description
- After grab sampling M=480 (8 hours of 60 minutes) consecutive periods on a shift with a non-Lognormal exposure distribution, the confidence δ that Ceiling noncompliance probability is $A_{C>H}W1\%$ is:
- for k=0 -> D(480,0,0.01)D99.2%
- for kW1 -> D(480,1,0.01)D95.3%

using the cumulative binomial test in TRUE EPISTAT.
Using the Poisson distribution on these data (p is small, n is large), with the number of cases expected 480*.01=4.8 and the number of cases observed 1, results in $D_{Poisson}$=99.23%. As could be expected the Poisson value is similar to the cumulative binomial.

Contents - Index

Example 37 Compiance with a ceiing limit value (distribution free)

*.HYG file Description
LEIDEL63 5 short period exposures to Hydrogen sulphide were sampled and tested against a ceiling of 20 PPM (Leidel 1977 page 65).
With A(W)=95%, GM=13.9 PPM and GSD=1.12 the NIOSH noncompliance probability (screen 25) is $A_{C>H}$=0.00073.
Because all 5 sampled periods were below the ceiling,
Leidel (1977 page 65) calculates the chance that the 16-5=11 nonsampled periods in the population are also in compliance as $P_c=(.9993)^{11}=.992$. In this example the following methodological shortcomings can be stipulated:

- GSD is (almost) completely determined by CV_t (see Example 57),
- The Normal shape is superior over the Lognormal,
- The standard normal method is biased for small sample size.

However, because the GSD is extreme small and overestimated, the chance on exceeding the ceiling value in the 11 nonsampled periods of this work situation, can be neglected.

- A M=25 grab sample, exposure sampling plan from a complete Lognormal distribution results in GM=1 and GSD=2.71818. With limit H=10 the NIOSH noncompliance probability is calculated as $A_{C>H}$/100=0.01065 indicating that k=1 out of M=93 is in noncompliance. The confidence that ceiling compliance is $\alpha_{C>H}$W5%, is estimated as D(93,1,0.05)=95%.Using the Poisson distribution on these data with the number of cases expected .05*93=4.65 and the number of cases observed 1, results in D(4.65,1)=94.6%. As could be expected the Poisson value is similar to the cumulative binomial. To get compliance with the Poisson test the sampling plan should be increased to at least 95 grab samples.

Contents - Index

Example 38 Compliance with a ceiling limit vaue (Noncentral Student)

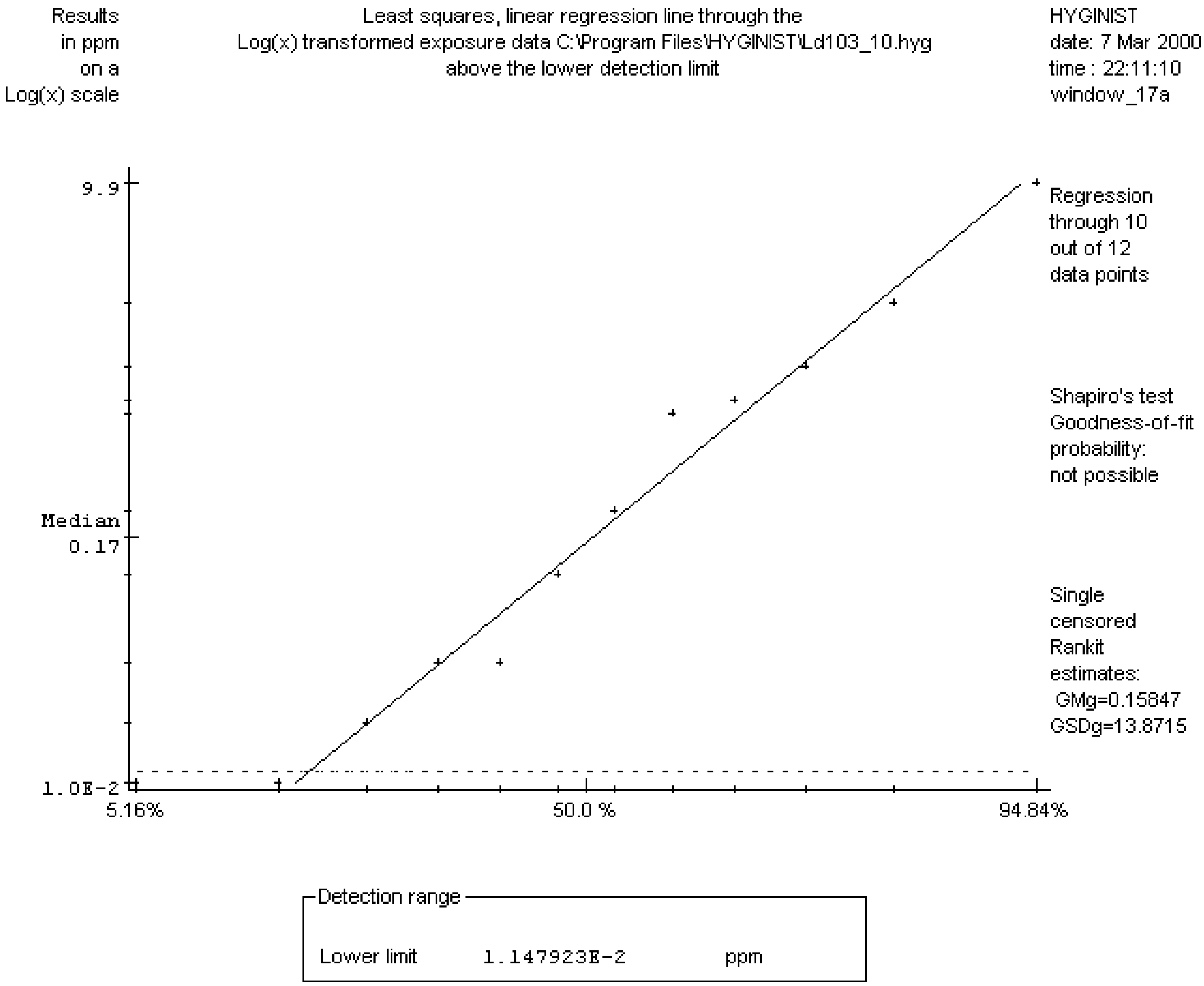
*.HYG file Description
LD103_10 The unbiased probability that a grab sample Hydrogen fluoride concentration (Leidel 1977 pages 103-104) is above the ceiling limit of 3 PPM HF, based on M`=10, LL=.01 PPM, GMg=.258 PPM and GSDg=13.87, is AC>H=19.9%. The confidence δ that Ceiling H is exceeded α<1% of the times is estimated as D(Cα<1%>H)=7.89% using the Noncentral Student (method 6 table 11). Since δ should be more than 95%, control measures are necessary to get in compliance with the ceiling.

LEIDEL63 5 TWAshort period Hydrogen sulphide were sampled and tested against a ceiling of H=20 PPM (Leidel 1977 page 65). With M=5, GM=13.9 PPM and GSD=1.12 the confidence δ that Ceiling H is exceeded α<5% of the times is estimated as D(Cα<5%>H)=87.7%, indicating noncompliance with the ceiling. Because the population is confined (see Example 54) recalculating the confidence with df=5 is permitted and results in D(Cα<5%>H)=91%. Since, however, the sample variance is completely determined by CVt the compliance conclusion can be drawn solely based on GM<H.

- Using an imaginary, grab sample, exposure sampling plan with GM=1, GSD=2.71828 and Ceiling H=100. If ceiling compliance is defined as 100-δ=αW1%, then MD13 grab samples are sufficient: D(Cα<1%>H)=99.2%

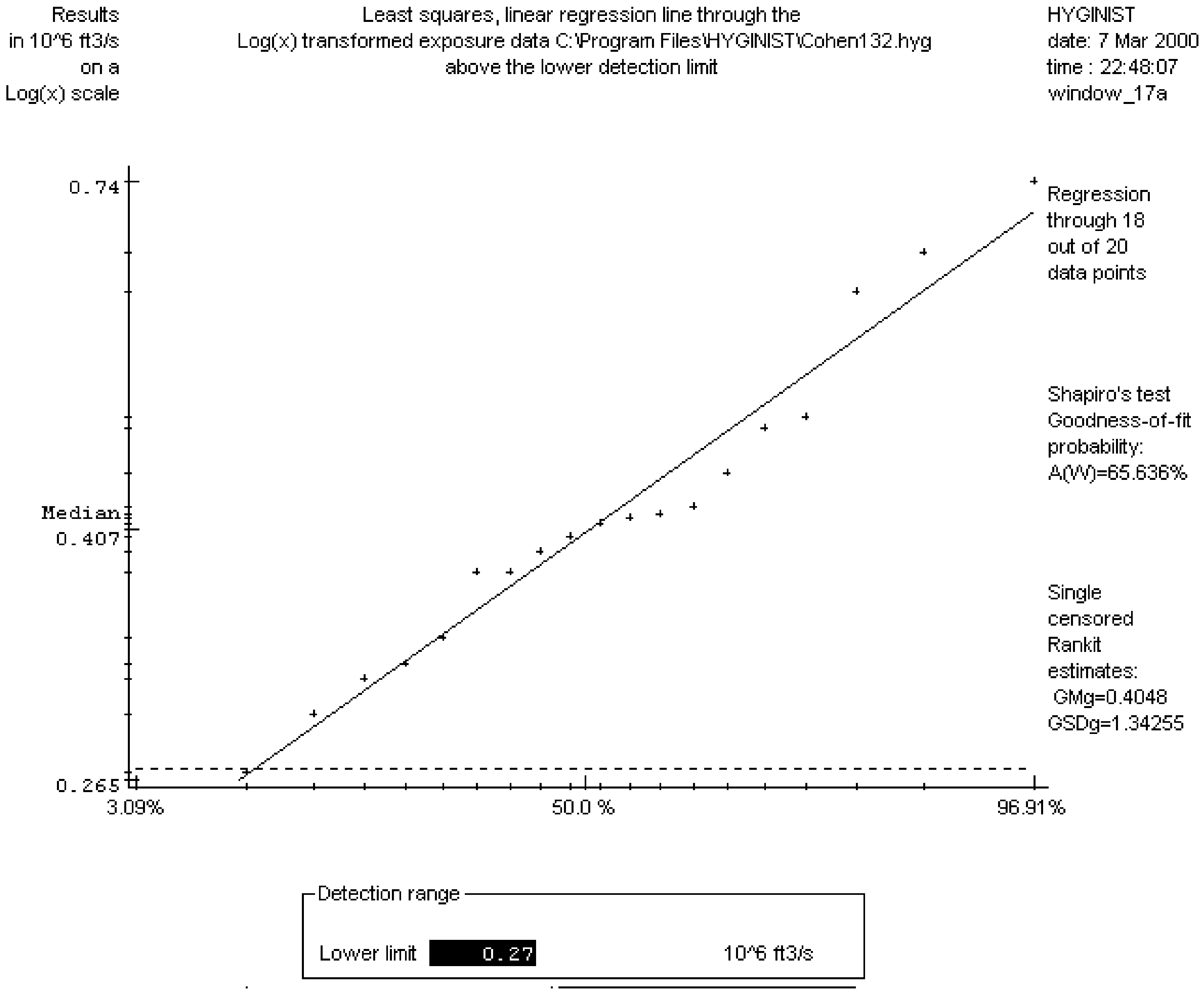
Contents - Index

Figure 13 Susquechanna river maximum flood levels in ft3/sec (COHEN132)



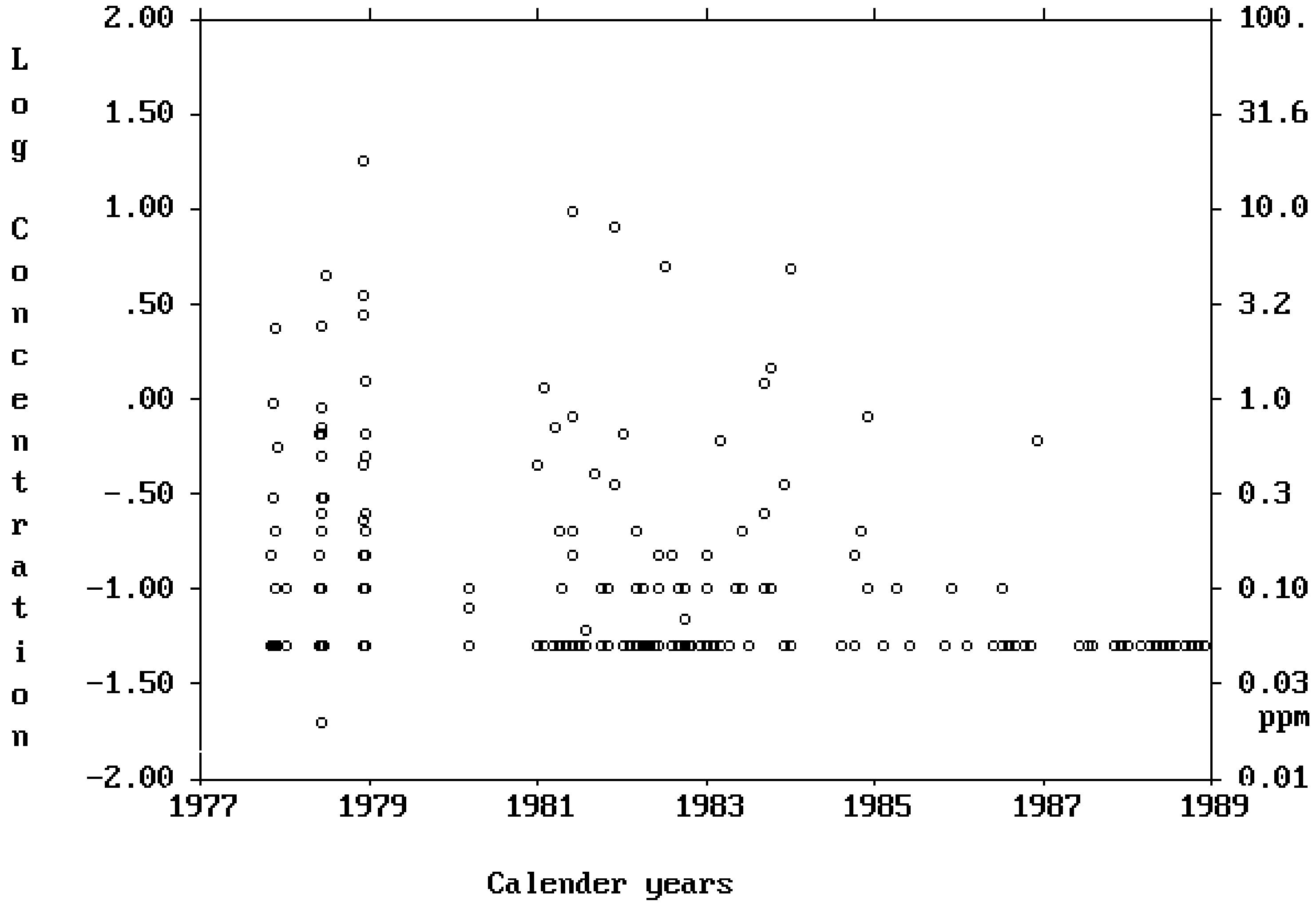
Contents - Index

Figure 14 12 CHF grab samples corrected for .1 PPM fixed background (LD103_10)



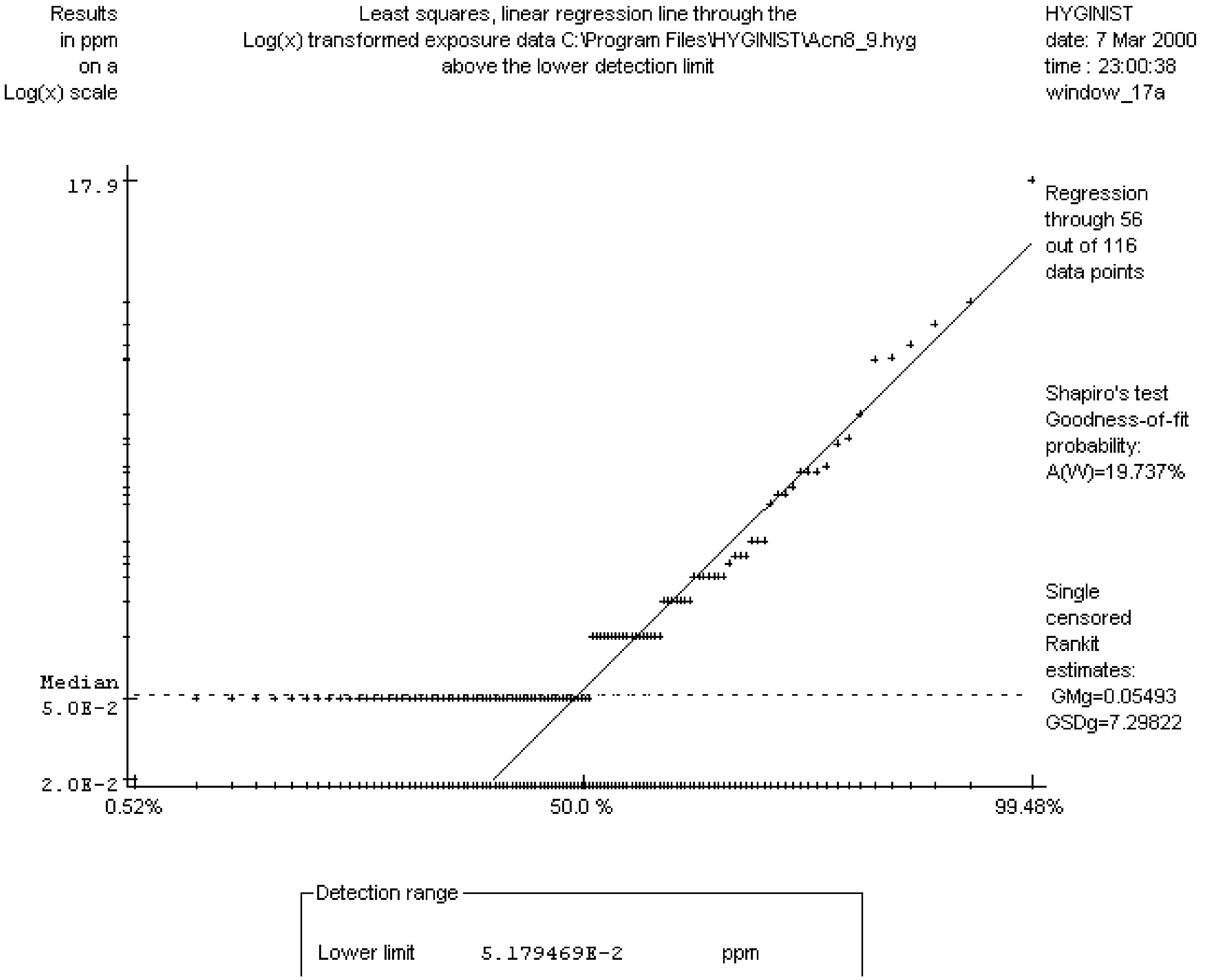
Contents - Index

Figure 17 283 TWA8 hour ACN in a SEG production



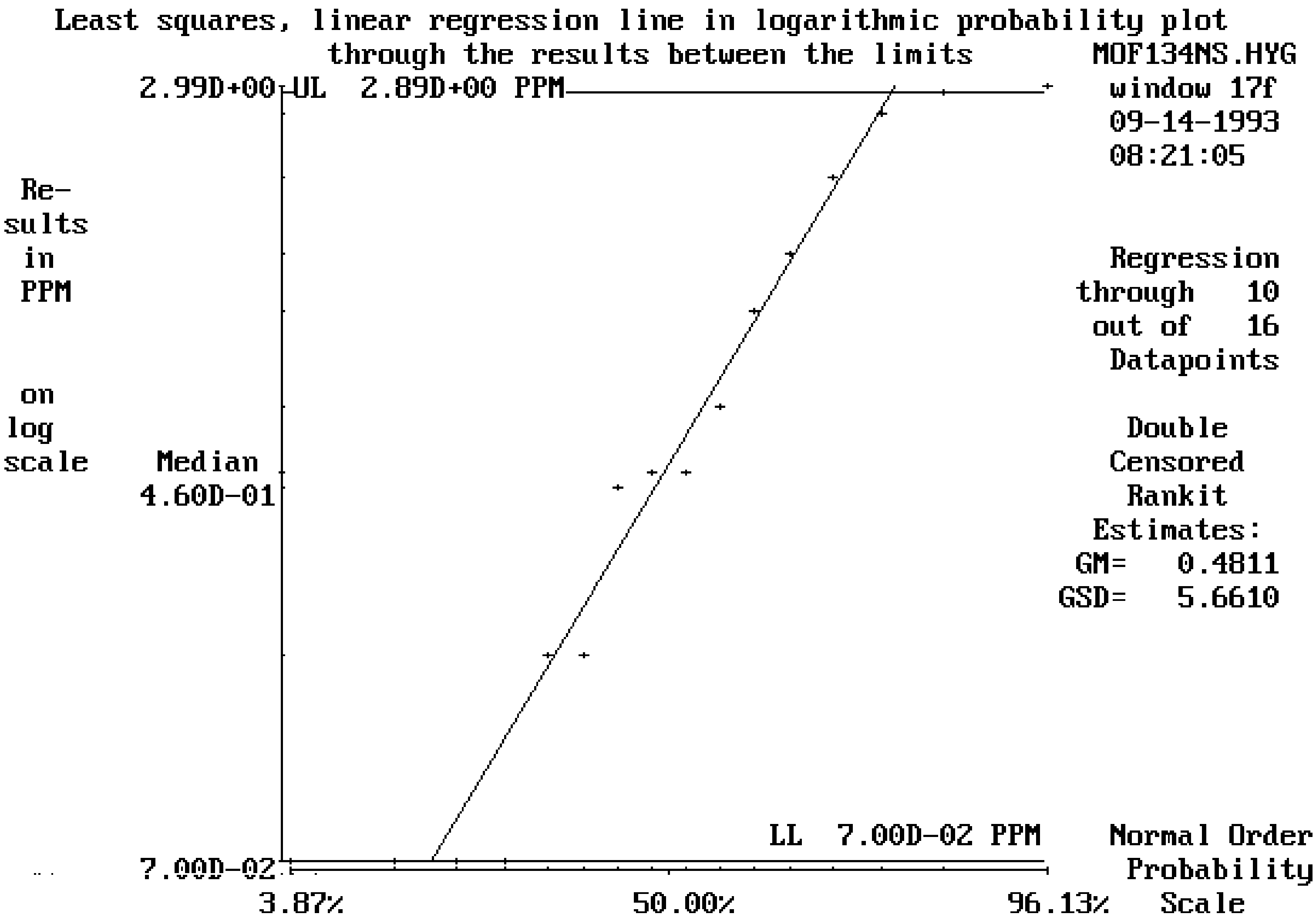
Contents - Index

Figure 18 116 TWA8 hour ACN in a SEG production 1977-78



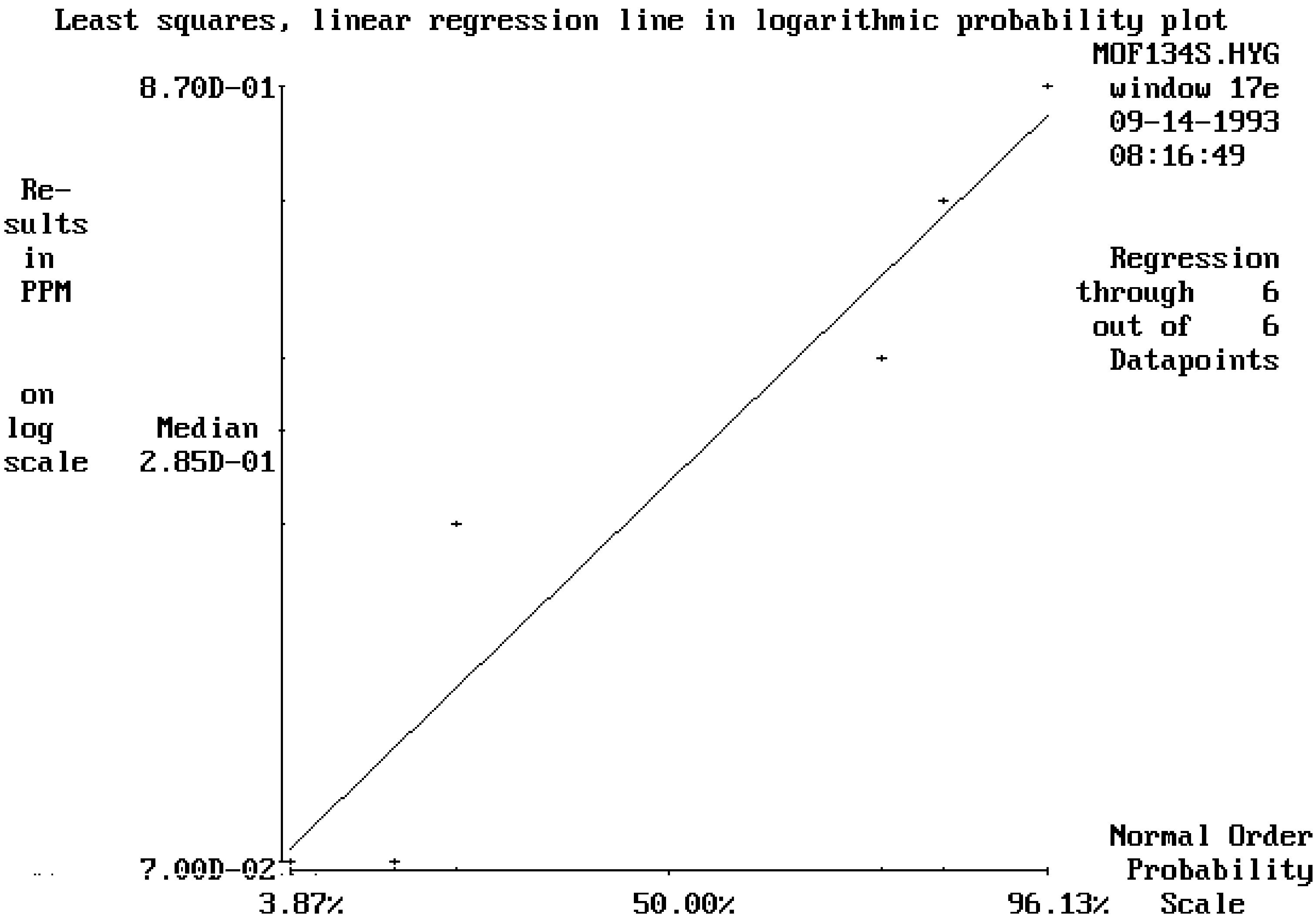
Contents - Index

Figure 19 16 TWA10-30 min Methoxyfluane. Nonscavenged results (MOF134NS)



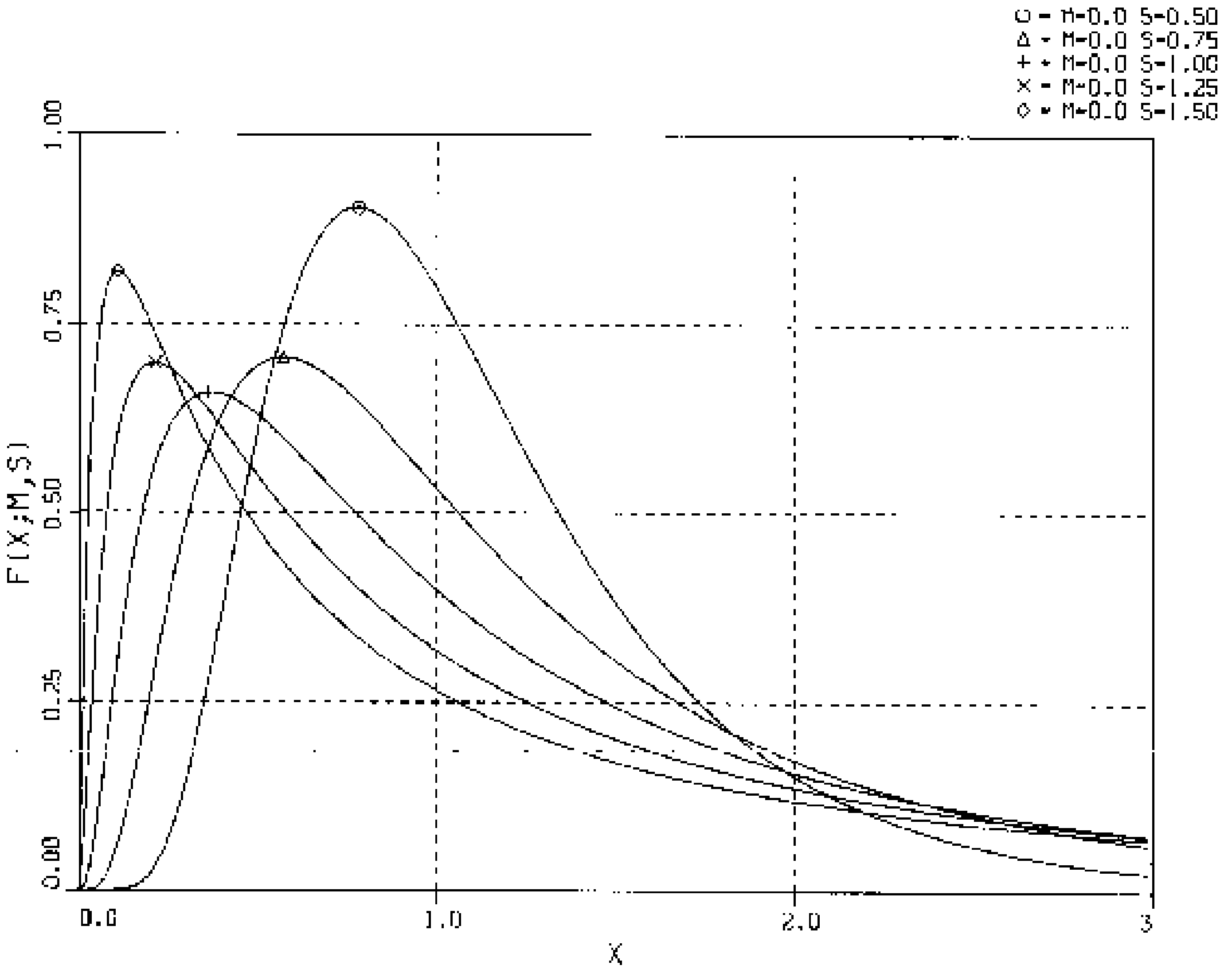
Contents - Index

Figure 20 6 TWA10-30 min Methoxyfluane. Scavenged results (MOF134S)



Contents - Index

Figure 21 A family of Lognormal distributions with mean M and standard deviation s

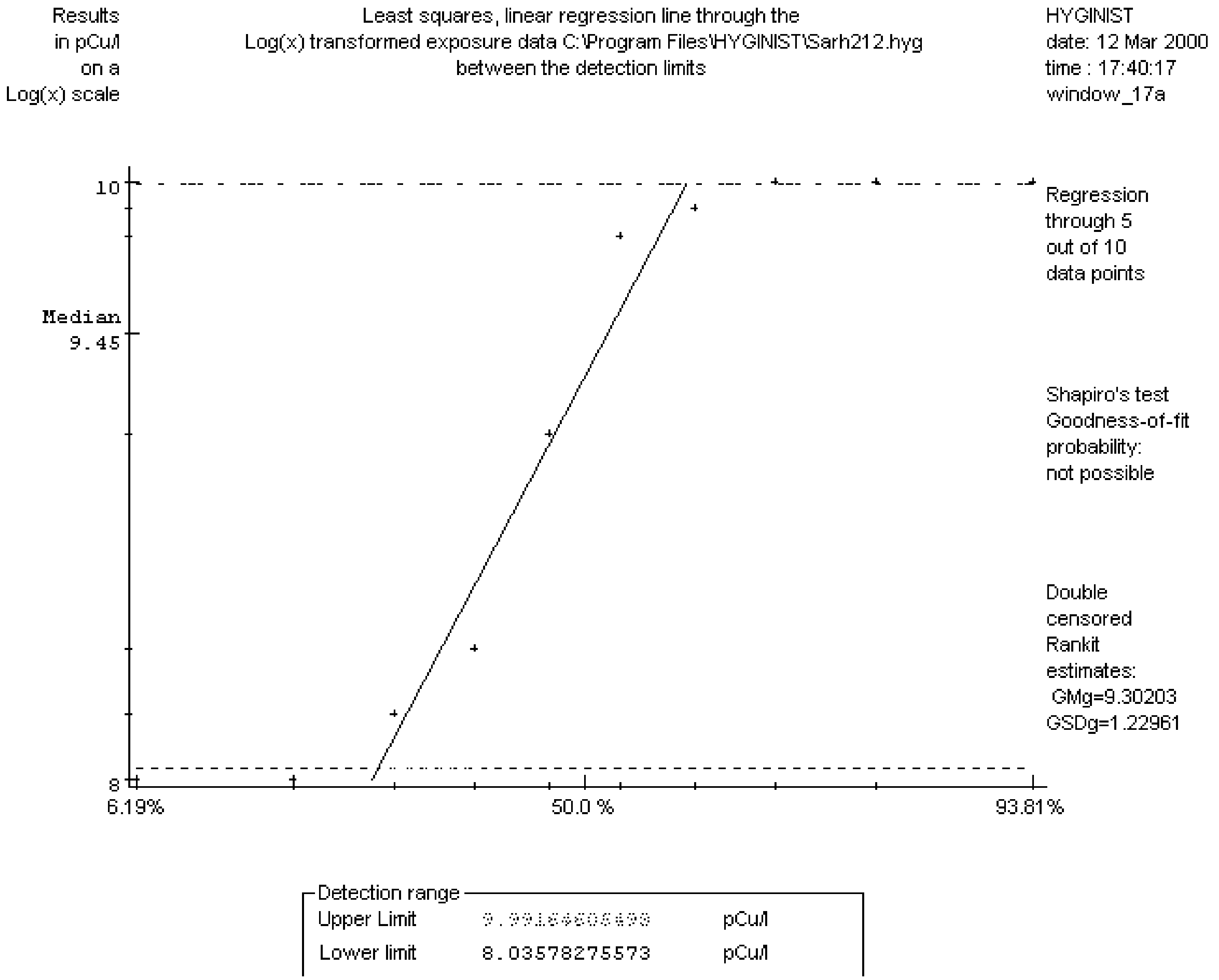


GM=exp(M)
GSD=exp(s)

M=0; GM=1
s=0.25; GSD=1.284
s=0.50; GSD=1.649
s=0.75; GSD=2.117
s=1.00; GSD=2.718
s=1.25; GSD=3.490
s=1.50; GSD=4.482

Contents - Index

Figure 26 10 measurements of Strontium-90 in milk (SARH212)



Contents - Index

Figure 27 Mice survival time (GUPTA271, Schneider 1986 p 69 & 88)