hyginist

☎ direct 31-46-765013 (GMT 14.00-16.00 hours)
Fax 31-46-752525
☎ home 31-43-622107 (19.00-21.00 hours)

Theo M.L. Scheffers, Author of HYGINIST computer program,
Cramer van Brienenstr 1F, 6225 BA Maastricht, The Netherlands
Bank: Postbank Arnhem, The Netherlands (Swift code PSTBNL2A) account number 3484314

---

homas Sheahan
Director of Communications
American Industrial Hygiene Association
P.O. Box 8390 345 White Pond Dr.
Akron, OH 44320 USA

---

your ref. your letter our ref. Maastricht, 11-09-91

re. HYGINIST Computer program

Dear mr Sheahan,

Enclosed the user's manual of the HYGINIST Computer program. As we agreed I have added a table of conntent, an introduction, a section on hard disk installation, an index and a flow chart. You can sent the contracts on the publication rights on the terms we agreed. Together with the signed contracts I will sent you the floppy disks with the WordPerfect Files, the HYGINIST program, the example files, the BASIC source listing and the dBase file with the names of those interested in HYGINIST.

Sincerely

Theo M.L. Scheffers
Author of HYGINIST computer
program and manual.

cc: Lex Burdorf
DOHS communications director

MANUAL. September 11, 1991

hyginist

A COMPUTER PROGRAM FOR THE QUANTITATIVE EVALUATION OF INDUSTRIAL HYGIENE EXPOSURE DATA

Theo M.L. Scheffers

Cramer van Brienenstr 1F
6225 BA Maastricht The Netherlands

Colophon

HYGINIST program and manual are made by Theo Scheffers.
Manual and program are distributed by the American Industrial Hygiene Association

Thanks to the DOHS members Jan Boleij and Eltjo Buringh, to Martha Waters and other members of the Statistical Tools Subcommittee of AIHA's Exposure Assessment Strategies Committee and to Gerald Cooper (Rhone Poulenc). Program and manual are the result of an uncontrollable passion (sometimes called avocation) for industrial hygiene, computers, statistics and wordprocessing. This result was not possible without the help of the DSM colleagues Wil ten Berge, Tom Hillegers and Ton Soons.

HYGINIST software is made with the highest possible accuracy. The author accept no liability for the misuse of the program. Users are invited to sent comments or suggestions that can improve the performance of the computer program to the author. Correspondents with accepted ideas will be mentioned in next versions of the manual.

SUMMARY

A COMPUTER PROGRAM FOR THE LOGNORMAL EVALUATION OF INDUSTRIAL HYGIENE EXPOSURE DATA

When industrial hygiene sampling periods form only a fraction of:
- the hygienic limit reference TWA period,
- the total employee exposure duration,

then statistical extrapolation becomes a important in exposure assessment. Because small samples and extrapolation to general results are common in Industrial Hygiene, unbiased statistical methods and estimators are preferred over more efficient but biased ones such as maximum likelihood.

The MS-DOS computer program HYGINIST Version 2.1 executes the extrapolation to unsampled periods of a series, mutually independent, measurement results of an industrial hygiene sampling plan to the unsampled periods in that specific exposure situation can be divided in three steps:
(1) Examining the Lognormal Goodness-of-fit,
(2) Estimating the Lognormal descriptive statistics,
(3) Extrapolation to unsampled periods and hypothesis testing.
Adequate and unbiased methods are available for all three steps.

(1) Plotting the cumulative frequency distribution of at least three, logarithmic transformed, results against the normal order probabilities (Teichroew 1953, Royston 1982b) is the best optical way to examine the conformity with the Lognormal model for both complete and censored samples. The analysis of variance W-test for normality (Shapiro 1965, Royston 1982c) is the most accurate among the omnibus tests for complete samples (Shapiro 1972, Filiben or d'Agostino 1971), especially when sample size is small.

(2) The antilogs of the mean and the standard deviation (GM and GSD respectively) of the logarithmically transformed results are the best estimators of the Lognormal descriptive statistics EXP(μ) and EXP(σ) for complete samples. If at least one result lies outside and three results inside the detection range, then GM and GSD can be derived from the least squares, linear regression line through the data points within the detection range, using the algorithms for the Type I, single and double sided, censored sample (Gupta 1952, Prescott 1970).

(3) If the exposure data refer to a series grab or partial period measurements which forms only a fraction of the reference period, GM and GSD are used for estimating the full period Time Weighted Average (TWA) concentration and for compliance testing with the Hygienic limit using:
- the uniform most powerful method (Bahr-Shalom 1976) for sample size odd and less than 25, and GSD smaller than 3.16
- the unrestricted approximate method (Jahr 1987).

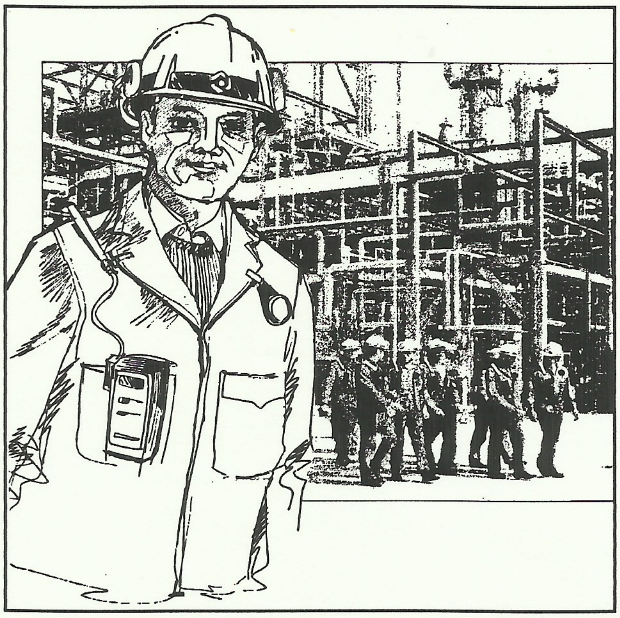
The unrestricted method is also of use for calculating the confidence interval of the average daily dose of a cumulative agent in occupational epidemiology.

If the exposure data refer to a series full reference period TWA measurements which forms only a fraction of the employee exposure duration of a homogeneous exposure situation, the long-term Noncompliance probability and the upper tolerance limit of the exposure is estimated using:
- the NIOSH method (Leidel 1977),
- the unbiased method (Wilks 1941) and
- the non-central Student method (Tuggle 1982).

Wilks and Tuggle estimate noncompliance with higher confidence than Leidel. The unbiased method is developed for quality control and is especially useful in minimizing the sample size for long-term compliance control.

Contents

INTRODUCTION



When your industrial hygiene sampling periods form only a fraction from:
- the hygienic limit reference TWA period or
- from the total employee exposure duration

then statistical extrapolation becomes important in exposure assessment.

The enclosed disk contains a MS-DOS computer program that evaluates the log-Normal shape of samples industrial hygiene exposure data and performs an up-to-date extrapolation to unsampled periods.

The growing use of the microcomputer in industrial hygiene offices offers now the opportunity to introduce this new generation, easy-to-use computer program that improves and facilitates exposure assessment and other evaluations on industrial hygiene sampling plans.

HYGINIST was first presented in the U.S. at the American Industrial Hygiene Conference on May 17, 1990 in Orlando, Florida. From the positive reactions at the AIHConference it was decided to develop an International version. The program was technically evaluated by the Statistical Tools Subcommittee from the Exposure Assessment Strategy Committee. The Committee advised AIHA to publish this program.

The eight characters of HYGINIST represents the MS-DOS command for starting the program on a personal computer. HYGINIST claims to be user friendly: with basic knowledge of the keyboard functions you must be able to install, start and operate the program. HYGINIST is developed to evaluate the log-Normal shape of industrial hygiene exposure data and for extrapolation and inference. It offers techniques which are not available in the current statistical computer programs.

☞ **The use of HYGINIST is only credible if the user is cognizant of industrial hygiene exposure assessment,** principally with occupational sampling strategy and compliance control (e.g. Leidel 1977). **Without basic knowledge of industrial hygiene, the use of HYGINIST is reprehensible.**

This manual describes the used statistics and their relation with exposure assessment. The manual is multi-functional: (1) It is a guide for operating the program, (2) It is a simplified workbook for the evaluation and extrapolation of industrial hygiene exposure data (3) and it describes the choices that were made in developing this program. The disk includes different example exposure data files from the industrial hygiene literature. The examples are used throughout the text of the manual. The Industrial Hygienist who want to know more about the used techniques, finds a comprehensive reference list with the original scientific literature.

11 September 1991

Theo Scheffers

1. RUNNING THE PROGRAM

With basic knowledge of the XT, AT or PS2 machinetype keyboard functions you must be able to install, start and operate the program. The eight characters of HYGINIST represents the MS-DOS command for starting the program. The program starts, if you enter:

HYGINIST

from the diskette drive prompt. With this command the IBM Color Graphics Adapter (CGA) is supported. A Graphics Adapter and Monitor is necessary to display the cumulative distribution plot (screen 17). Other IBM Graphics Adaptors are supported with a higher resolution than GCA. In the table the start command from the prompt are mentioned for the different types of IMB Graphics Adapter.

Start Command from the prompt A:\>	Graphics adapter type
HYGINIST /10	Enhanced Graphics Adapter (EGA),
HYGINIST /11	Multi Color Graphics Array (MCGA),
HYGINIST /12	Video Graphics Array (VGA),

Non-IBM type of color adaptor are not supported. In combination with a monochrome monitor it is possible that a Hercules Graphics Card, Hercules Graphics Card Plus, Hercules InColor adapter, Olivetti Color Adapter or AT&T Adapter Boards display the cumulative distribution are supported if you enter:

HYGINIST /04

Try before you buy. Official owners can call for emergency assistance to Theo Scheffers in the Netherlands 31-46-765013 (office, 14-16 hrs GMT) or 31-43-622107 (home 19-21 hrs GMT).

1.1. Windows

The 34 windows of the program are numbered in the upper right corner of the screen, where you also find the time and date. You can copy a window to a file if the message **<F>=ScrD** is indicated in the lower right corner (Windows 15, 18, 20, 21, 24 through 27 & 29 through 34). The ASCII-file is saved in the current directory as **<data file name>.<window number>** except window 18a which is saved as **<data file name>.<transformation>**. The copy can be read by almost any wordprocessing program. If you have a graphics screen capture-program (like WordPerfect's GRAB) you can capture window 17 and use it in reports.

1.2. Window handling

The handling of the program is made very easily:

- Windows without (multiple) choices (Windows 1ac, 3, 8, 11, 17ac and 20) are automatically advanced after some reading time or by pressing **<any key>**.
- The **<Esc>** key will take you back one or more windows in the program.
- Window 1b includes the license agreement. If you press the **<Enter>** the you accept the license agreement. Any other key will return to your system prompt.

- If the window ask you to make a choice (Windows 2 , 4, 5, 17bcdf 21, 22, 23, 28ab, 29c) then press the desired number or Capital character to continue. Using the **<Esc>** key will take you back one step in the program.
- You can stop the running program by pressing the **<Ctrl>** and **<Break>** keys at the same time.

1.3. Entering characters or numeric values.

Units of measurement (window 3b) and the data file name (windows 4, 9 and 12) comprise of at most 8 alfa-numeric characters. A non-integer numeric value (window 6 or 15) comprises of at most 16 digits for the mantissa, and **<D>** or **<E>**, **<+>** or **<->** and 1 or 2 digits for the exponent. Type digits and characters and press **<Enter>**. Avoid the use of the comma (a comma will give you a "Redo from start ?" message). You can correct characters or numbers with the **<Delete>** or the **<Backspace>** key.

2. DATA ENTRY

HYGINIST accepts exposure data as:
- A series of 2 through 2000 industrial hygiene measurements from a disk (windows 8 and 9) or from the keyboard (windows 3 through 7).
- The Lognormal descriptive statistics GM and GSD from the keyboard.

Exposure data should meet the following requirements:
- sample size between 2 and 2000,
- Geometric mean GM $>$1D-10 and $<$1D+10
- Geometric standard deviation GSD $>$1 and $<$10000.
- individual results must have a value of $\geq$1D-10 and $\leq$1D+10.

HYGINIST imports a set results in ASCII DELIMITED or SDF format, if:
- the file contains only the numeric values of 2 to 2000 measurement results,
- the results are separated with comma's or **<Carriage return line feed>**,
- the file name should conform the Standard DOS file name restrictions
- the obligate name extension is **<.HYG>**.

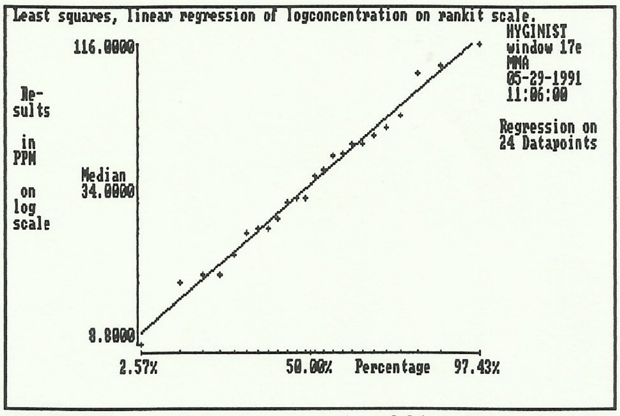
The program diskette contains 34 files with exposure data which are from Indutrial Hygiene literature.

Results are automatically sorted (window 10) in ascending numerical order using an adapted Shell-Mezgar algorithm (Press 1987 p288). HYGINIST files are import compatible with standard computer programs like dBase, Lotus, WordPerfect, SPSSPC and PCTOOLS.

3. GOODNESS OF FIT

According to Esmen (1977) the Lognormal distribution is the best estimator of a mutually, independent population of industrial hygiene results. 'Best' means unbiased or consistent, with minimum variance (window 16).
If the Goodness-of-fit procedure excludes a Lognormal distribution, then stop HYGINIST quantitative evaluation and examine:

- the population base of the sample,
- the sampling strategy,
- the analytical method.



HYGINIST's cumulative frequency plot of 24 methyl methacrylate averages (Leidel 1977 p105).

3.1. Lognormal probability plot

HYGINIST plots the cumulative frequency distribution on expected normal order statistics scale (window 17a). Expected Normal order statistics (= Rankits) are the unbiased, minimum variance estimates of the deviates of the standard normal distribution (Teichroew 1956 & Royston 1982a). They are equivalent with the plotting positions for normal probability paper (Leidel 1977 p100-101).

3.2. Least squares, linear regression

The least squares, linear regression line through the plotting positions (window 17e), is calculated with the algorithms 9.2.3 and 9.2.4. of Snedecor (1980).

3.3. The omnibus W-test for normality

If the sample is complete (uncensored), then the one-sided probability for normal shape is calculated with the omnibus W-test (Shapiro 1965), using the log-transformed results and Algorithm AS 181 (Royston 1982b). The power to detect real deviation from normality is poor, if sample size is small. The W-test has the highest accuracy among the omnibus tests (like Filiben 1975 or d'Agostino 1971). An omnibus test for shape for a censored sample is not included (window 18b).

4. DESCRIPTIVE STATISTICS

GM, GSD and AM are estimators of the true geometric mean EXP(μ), geometric standard deviation EXP(σ) and the arithmetic mean ß, respectively (window 20). If a Hygienic limit (H) is entered (window 21a option 2) the number and percentage of results below the limit are computed. The desired percentage (window 21 option 1) is used to calculate upper confidence and tolerance limits in exposure assessment. The use of the default U=95% is common practice in industrial hygiene (Leidel 1977 p69 and 118).

4.1. Complete sample

If the sample is uncensored GM and GSD are calculated (window 20a) as the antilogs of the mean and the standard deviation of the logarithms of M results, respectively:

$$x_i = \mathrm{LOG}(c_i) \tag{1}$$

$$æ = 1/M * \Sigma x_i \tag{2}$$

$$GM = \mathrm{EXP}(æ) \tag{3}$$

$$s = \{\Sigma(x_i - æ)^2/(M-1)\}^{+1/2} \tag{4}$$

$$GSD = \mathrm{EXP}(s) \tag{5}$$

The unbiased estimator of the arithmetic mean ß with the highest efficiency is:

$$AM = GM * \Phi(s^2/2) \tag{6}$$

with $\Phi(s^2/2)$ a Bessel type function:

$$\Phi(t) = 1+(M-1)/M*t+(M-1)^3/[M^2*(M+1)]*t^2/2!+ (M-1)^5/[M^3(M+1)(M+3)]*t^3/3!+.......... \tag{7}$$

Oldham (1953) mentions formula (6) and (7) in his industrial hygiene citation classic and refers to Finney (1941). The algorithm converges in 7 or 8 terms to a relative accuracy of less then 10^{-8}.

4.2. Censored sample

If at least one result lies outside the detection range (window 17bcd), then GM and GSD are derived from the least squares, linear regression line through the data points within the detection range, using the algorithms for the Type I, single and double sided, censored sample (Gupta 1952, Prescott 1970). For **J** (j=1,...J) results within the detectionrange, GSD is calculated using:

$$GSD^g = \mathrm{EXP}((\Sigma(R_j x_j) - \Sigma R_j \Sigma x_j)/(\Sigma(R_j)^2 - (\Sigma R_j)^2/J))$$

The geometric mean GM^g is estimated by interpolation to the median:

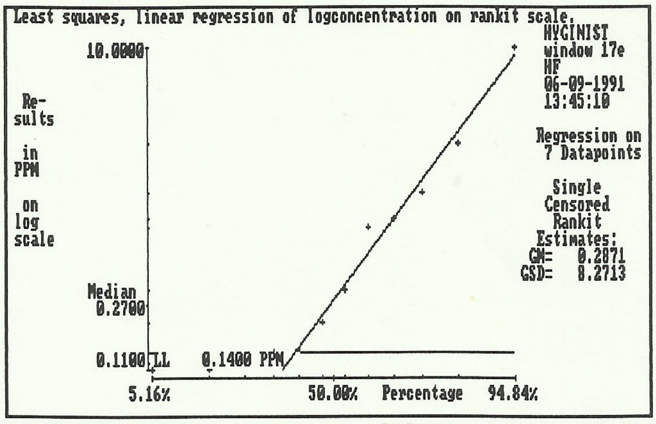
$$GM^g = \mathrm{EXP}\{(\Sigma x_j)/J - \log(GSD^g) * \Sigma R_j/J\} \tag{9}$$

The linear estimates of Gupta (1952) for type II censered samples are equivalent with formula (8) and (9). HYGINIST calculates AM^g from $GSDG^g$, GM^g, the sample size J from within the detection range and formula (7).

5. EXPOSURE ASSESSMENT

The used method of extrapolation depends on whether your industrial hygiene sampling periods form a fraction of the hygienic limit reference period (5.1.) or form a fraction of the total employee, exposure duration (5.2.).

5.1. Classification of a full period exposure.



HYGINIST's cumulative frequency plot of 12 HF concentrations (Leidel 1977 p103).

If the exposure data refer to a series grab or partial period measurements which forms only a fraction of the reference period, GM and GSD are used for estimating the full period TWA concentration and for compliance testing with the Hygienic limit using:

- the uniform most powerful compliance test (Bahr-Shalom 1976), for sample size odd, $M<25$, and $GSD<1.65$
- an (unrestricted) approximate compliance test (Jahr 1987) with test statistic:

$$t= LOG(AM/H)*SQR(df)/LOG(GSD) \quad (11)$$

where t follows the student distribution with $df=M-1$ (with M the data within the detection range). Compliance probability is calculated with the algorithms of Owen (1968 p465).

The following two data sets refer to extrapolation to a full Hygienic limit reference period.

During one shift, an employee's exposure was sampled with 8 charcoal tubes. Each tube was exposed for 20 minutes. From these data the TWA exposure over the full shift is estimated. Compliance with the standard is estimated using the procedure for (random) grab samples (Leidel 1977 §4.2.3). If the values of the arithmetic mean of logarithm values y=10log(GM/H) and the standard deviation of the logarithms s=10log(GSD) are entered in the nomograms of Bahr-Shalom (1975) the exposure is classified as possible overexposure (uniform most powerful test). The HYGINIST results (window 24) are in close agreement with the outcome of Leidel (1977 p.57). Note that the untransformed data give a better Goodness-of-fit with the Normal distribution than the logarithmic transformed results. A possible explanation is that the influence of CV_t *on the variance is higher than the 0.06 from the table in Appendix D (Leidel 1977 p.80).*

```
   INDUSTRIAL HYGIENE EXPOSURE DATA EVALUATION AND EXTRAPOLATION  HYGINIST
   ************************D E M O  disk************************  window 24
                                                                  05-29-1991
  Test for compliance (Non full period sample)                    11:23:18
  Sample size of ETHYLALC                          M=     8
  Geometric mean EXP(μ)                           GM= 1004.5512      PPM
  Geometric standard deviation EXP(σ)            GSD=    1.3805
  Arithmetic mean ß                               AM= 1051.0722      PPM
  Hygienic limit value                             H= 1000.0000      PPM

  Noncompliance probability α               A(ß≥H)=   65.2501        %
  Upper confidence limit of ß        AM(δ≥95.0000%)= 1324.1014       PPM
  Mean of log(relative results)        10log(GM/H)=     1.9721D-03
  Standard Deviation of log(results)    10log(GSD)=     0.1401
  Test on non full-reference-period samples (random grab or partial period,
  Leidel 1977 p38 & 60), using Jahr's Algorithm 11 (Staub 47 (1987) 153-156)
  For the 'uniformly most powerful' test, use 10log(GM/H)=y, 10log(GSD)=s &
  fig 1&2 in Bar-Shalom (AIHAJ 37 (1976) p472). Noncompliance if A(ß≥H)>95%,
  compliance if A<5% and possible overexposure in between (Leidel 1977 p58).
```

Hydrogen Fluoride exposure near a control panel in a production unit was measured using a stationary sequential sampler. Because of a fixed low background level of 0.11 PPM, the cumulative frequency distribution lacked lognormality in the left tail (see figure page 9). Leidel (1977 p.103) encountered this mixed emission using the three parameter Lognormal model and found for the random variation exposure GM=0.26 PPM and GSD=12.8. HYGINIST excludes constant background levels by setting a lower imit (0.14 PPM). From the linear regression line through the 7 result above the lower limit, the random variation exposure is estimated as GM=.029 PPM and GSD=8.27. The arithmetic mean for the random variation exposure is 1.424 PPM The full period TWA is 1.424+0.11=1.53 PPM (window 24).

```
   INDUSTRIAL HYGIENE EXPOSURE DATA EVALUATION AND EXTRAPOLATION  HYGINIST
   ************************D E M O  disk************************  window 24
                                                                  05-29-1991
  Test for compliance (Non full period sample)                    11:27:24
  Results between detection limits of HF          M'=     7
  Geometric mean EXP(μ)                           GM=     0.2871     PPM
  Geometric standard deviation EXP(σ)            GSD=     8.2713
  Arithmetic mean ß                               AM=     1.4240     PPM
  Hygienic limit value                             H=     3.0000     PPM
  Noncompliance probability α                A(ß≥H)=    21.0422       %
  Upper confidence limit of ß        AM(δ≥95.0000%)=     7.6107     PPM
  Mean of log(relative results)        10log(GM/H)=    -1.0191
  Standard Deviation of log(results)    10log(GSD)=     0.9176
```

5.2. Long-term compliance control

If the exposure data refer to a series, mutually independent, full reference period TWA measurements which forms only a fraction of the employee exposure duration of a homogeneous exposure situation, the long-term Noncompliance probability and the upper tolerance limit of the exposure is estimated using:

- the NIOSH method (Leidel 1977),
- the unbiased estimate (Wilks 1941) and
- the non-central Student method (Tuggle 1982).

The test statistic for the NIOSH method (window 25) is:

$$Z = LOG(H/GM)/LOG(GSD) \quad (8)$$

where Z follows the standard normal distribution. Compliance probability is calculated using the algorithm of Abramowitz (1970 § 26.2.11). The industrial hygiene situation is considered non-compliance, if the probability is more than 5% (Leidel 1977 p69). Leidel's method overestimated compliance probability on the average. However, confidence $D(A>\alpha)$ of the test statistic is at most 50% and decreases with decreasing sample size M and/or increasing GSD!

The test statistic for the unbiased estimate (window 26) of compliance probability (Wilks 1941, proven by Proschan 1953) is:

$$t = LOG(H/GM)/LOG(GSD)/SQR(1+1/M) \quad (9)$$

where t follows the central Student distribution with df=M-1 the number of degrees of freedom. One-sided probability is calculated with the algorithms of Owen (1968 p465). Confidence $D(A>\alpha)$ is at least 50% and **increases** with decreasing sample size M and increasing GSD.

Confidence of a desired compliance probability (=Tolerance) is introduced by Tuggle (1982). The test statistic (window 27) is:

$$T_{df,\delta} = LOG(H/GM)/LOG(GSD) \quad (10)$$

where T follows the Non-central Student distribution with df=M-1 and δ the standard normal deviate of the desired compliance probability. One-sided confidence is calculated using the algorithms of Owen (1968 p463-465).

The following two data sets refer to long-term extrapolation:

In the job category "Mix man" in a facility using Methyl Methacrylate TWA concentrations were measured on 24 employees. Leidel's figure 1-5 (Leidel 1977 p105) is in close agreement with the rankits plot and regression line of window 17e of (see page 8 of this manual). The descriptive statistics GM and GSD are also in close agreement.

```
  INDUSTRIAL HYGIENE EXPOSURE DATA EVALUATION AND EXTRAPOLATION  HYGINIST
  ************************D E M O  disk************************  window 20
                                                                 05-29-1991
Exposure data MMA.HYG                                            13:52:44
Sample size                                       M=   24
Direct estimators of the descriptive statistics:
Geometric mean EXP(μ)                            GM=   34.5198      PPM
Geometric standard deviation EXP(σ)             GSD=    1.8862
Arithmetic mean ß                                AM=   41.8054      PPM
The direct estimators GM and GSD are the antilogs of the mean and the stan-
dard deviation of the logarithms of the complete sample.
AM is calculated from GM and GSD (Finney  JRSS suppl. 7 (1941) p155-161).
```

On 10 different days in a period of 6 months an employee's 8-hour TWA Dioxane (H=100 PPM) concentration was measured. The cumulative distribution showed close conformation with the Lognormal distribution (A(W)=99.5%). The long-term compliance probability equals the results of Leidel (1977 p68 and 69), both indicating that engineering controls should be implemented in this situation.

```
  INDUSTRIAL HYGIENE EXPOSURE DATA EVALUATION AND EXTRAPOLATION  HYGINIST
  ************************D E M O  disk************************  window 25
                                                                 05-29-1991
Long-term compliance control (NIOSH)                             13:53:47
Sample size of DIOXANE                            M=   10
Geometric mean EXP(μ)                            GM=   78.4309      PPM
Geometric standard deviation EXP(σ)             GSD=    1.6310
Arithmetic mean ß                                AM=   87.2574      PPM
Hygienic limit value                              H=  100.0000      PPM
Noncompliance probability α                 A(C≥H)=    30.9711       %
Upper tolerance limit                 C(Γ≥95.0000%)=  175.3562      PPM
A strong indication exists that engineering controls should be installed if
the noncompliance probability A>5% (Leidel 1977 p69). A(C≥H) overestimates
α, on average. Confidence δ(A≥α) however, is less than 50% and decreases,
with decreasing sample size M and/or increasing GSD.
```

6. CONFIDENCE AND TWO SAMPLE INFERENCE

HYGINIST compares the current exposure data with a set of reference data. Valid reference data are the true Lognormal descriptive statistics EXP(μ) and EXP(σ), or the sample estimates GM and GSD. Two sample inference, needs at least sample size M2 and GSD2. GSD's are compared first. To compare GM's both sample GSD's should be known. This transaction performs:

- inference statistics and confidence calculations,
- combining the descriptive statistics of two samples with the same population base.

6.1. Two sample GSD's

The squared ratio of the logarithms of two GSD's (window 29)

$$F= \{LOG(GSD)/LOG(GSD2)\}^2 \quad (12)$$

follows the Fisher distribution (Abramowitz 1970 §26.6.2). The two sided probability is calculated using the algorithms 26.6.4/5/8, or 26.6.15 for M*M2>32676 (Abramowitz 1970). HYGINIST lets you decide if it is likely that the GSD's are equal. This information is needed to compare the GM's.

6.2. Two sample GM's

The log(GM) difference of two Lognormal samples, adjusted for the combined variances, follows Student's t distribution. When GSD's are equal (window 30) the degrees of freedom are summed (Snedecor 1980 §6.9). Satterthwaite's approximate number of degrees of freedom (Snedecor 1980 §6.11) is used when GSD's are unequal (window 31).

6.3. True EXP(σ)

The standardized variance

$$X^2= (M-1)*\{LOG(GSD)/\sigma\}^2 \quad (13)$$

follows a Chi-Square distribution (Abramowitz 1970 § 26.4.2) with df=M-1. The two-sided probability that GSD equals σ is calculated using algorithms 26.4.4 & 26.6.5 or 26.4.14 of Abramowitz (1970) (window 32).

6.4. True EXP(μ)

The standardized difference of μ and Log(GM)

$$t= \{\mu-LOG(GM)\}*SQR(M)/LOG(GSD) \quad (14)$$

follows a Student distribution with df=M-1 (Snedecor 1980 §4.9). The two-sided probability is calculated with the algorithms of Owen (1968 p465) (window 33).

7. MINIMIZING SAMPLE SIZE

From GM, GSD, Hygienic Limit H and formula (9), the lowest sample size M2 that gives the unbiased estimate for the desired long-term compliance control probability is calculated (Wilks 1941).

HYGINIST will respond with one of the following statements:

- No such sample size exists, change the industrial hygiene situation.
- The sample size is more than 50, consider if changing the industrial hygiene situation is more profitable than performing a very large amount of industrial hygiene measurements.
- The sample size is at most 2. Consider if an industrial hygiene sampling program is necessary in this situation.
- The necessary sample size for the industrial hygiene sampling program is xx (a specific value between 2 and 50) and this is less/equal/more than the current sample size.

References

Abramowitz M., Stegun I.A. Handbook of mathematical functions. Dover Publications Inc. New York (1964 5ᵉ edition 1970)

d'Agostino R.B. An omnibus test of normality for moderate and large size samples. Biometrika 58 (1971) 341-348.

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.

Esmen N.A, Hammad Y.Y. Log-normality of environmental sampling data. J. Environ. Sci. Health. 162 (1977) 29-41.

Finney D.J. On the distribution of a variate whose logarithm is Normally distributed. J. Royal Statistical Society Suppl., 7 (1941) 155-161.

Filiben J.J. The probability plot correlation coefficient test for normality. Technometrics 17 (1975) 111-117.

Gupta A.K. Estimation of the mean and the standard deviation of a normal population from a censored sample. Biometrika 39 (1952) 260-273.

Jahr J. Calculations for workplace measurements and other data with a lognormal distribution. Staub. Reinhaltung der Luft. 47 (1987) 153-156.

Leidel N.A., Bush K.A., Lynch J.R. Occupational exposure sampling strategy manual. DHEW, NIOSH publ. 137, Cincinnati (1977).

Oldham P.D. The nature of the variability of dust concentration at the coal face. Brit. J. Ind. Med. 10 (1953) 227-234.

Owen D.B. A survey of properties and applications of the noncentral t-Distribution. Technometrics 10 (1968) 445-477.

Press W.H. et al Numerical Recipes. The art of scientific computing. Cambridge (1987) p229 en p724.

Prescott P. Estimation of the standard deviation of a normal population from double censored samples. Biometrika 57 (1970) 409-419.

Proschan F. Confidence and tolerance intervals for the normal distribution. Am. Statistical Assoc. J. 48 (1953) 550-564.

Royston J.P. Expected Normal order Statistics (exact and approximate). Applied Statistics 31 (1982a) 161-165.

Royston J.P. The W-test for normality. Applied Statistics 31 (1982b) 176-180.

Shapiro S.S., Wilk M.B. An analysis of variance test for normality. Biometrika 52 (1965) 591-611.

Snedecor G.W. Statistical methods (1980)

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.

Tuggle R.M. Assessment of occupational exposure using one-sided tolerance limits. Am. Ind. Hyg. Assoc. J 43 (1982) 338-346

Wilks S.S. Determination of sample sizes for setting tolerance limits. Ann of Mathematical Statistics 12 (1941) 91-96.