3. Practice – 2014

Confidence interval, Random error propagation

1. Exercise

The drying time of an oil paint was investigated. n = 150 observations were made, the average drying time was 1.25 hours. The standard deviation is known to be 0.41 hours.
Calculate the result of the measurement using the average and the confidence limits at the significance level of 98%!

Solution:

We need to calculate the radius of the confidence interval (a) around the mean ($\bar{x}$), so that the expected value ($M(\xi)$) falls into the interval with the probability p. Formulating the problem we can write:

$$P(\bar{x} - a \leq M(\xi) \leq \bar{x} + a) = p$$

We know that

i. the average is an unbiased estimation of the mean, that is $\boldsymbol{M(\bar{x})}$**=** $\boldsymbol{M(\xi)}$;
ii. **when** $\boldsymbol{D(\xi) = \sigma}$ **is known** $D(\bar{x}) = \frac{\sigma}{\sqrt{n}}$.

Hence, we rewrite the inequality in the brackets so that in the middle we have a standardized random variable:

$$P\left(-\frac{a}{\frac{\sigma}{\sqrt{n}}} \leq \frac{\bar{x} - M(\xi)}{\frac{\sigma}{\sqrt{n}}} \leq \frac{a}{\frac{\sigma}{\sqrt{n}}}\right) = p$$

Let $\eta = \frac{\bar{x} - M(\xi)}{\frac{\sigma}{\sqrt{n}}}$, and $\lambda = \frac{a\sqrt{n}}{\sigma}$. So

$$P(-\lambda \leq \eta \leq \lambda) = p$$

From the Central Limit Theorem we know that $\eta \in N(0,1)$ (η is normally distributed with mean 0 and standard deviation 1. Since the probability if belonging to an interval can be calculated from the cumulative density function of η which we denote by $\Phi(x)$:

$$p = P(-\lambda \leq \eta \leq \lambda) = \Phi(\lambda) - \Phi(-\lambda) = 2\Phi(\lambda) - 1,$$

Where by definition $\Phi(x) = P(\eta < x)$. The last equality comes from the fact that if η is standard normal distributed, then $\Phi(-x) = 1 - \Phi(x)$. Comparing the left and right hand side of the above equation, we get $\Phi(\lambda)$:

$$\Phi(\lambda) = \frac{p+1}{2} \;\Rightarrow\; \lambda = \Phi^{-1}\left(\frac{p+1}{2}\right)$$

$$\boldsymbol{a = \frac{\lambda\sigma}{\sqrt{n}}}$$

- Fill out the excel table:
  - Number of data: **D12:=150**
  - The mean of the data: **D13: =1.25**

- The standard deviation: **D14: =0.41**
- The significance level p: **D16:=0.98**

- To determine λ:
  - Calculate (p+1)/2: **D17:=(D16+1)/2**
  - Find λ for p: **D19:= NORM.S.INV(D17)**
- Calculate the confidence limit: $a = \lambda * \frac{\sigma}{\sqrt{n}}$
  - **D21:= D19*D14/SQRT(D12)**
- Then the answer is: The drying time of the oil paint is 1.25 ± 0.08 hours with the confidence level p=98%.

2. Exercise

The filling weights of a detergent were measured using a direct measurement method. The results of the measurement are included in the table below.

a.) Calculate the results of the measurement using the average and the confidence interval at a significance level of 95%!

b.) Let us assume that the calculated sample standard deviation is constant with the increasing numbers of the measurement. How many measurements are required to reduce the relative error below 2%?

Solution

When $D(\xi) = \sigma$ is unknown, s^* is an unbiased estimation of σ, hence we write:

$$P\left(-\frac{a}{\frac{s^*}{\sqrt{n}}} \leq \frac{\bar{x} - M(\xi)}{\frac{s^*}{\sqrt{n}}} \leq \frac{a}{\frac{s^*}{\sqrt{n}}}\right) = p$$

Now $\eta_{st} = \frac{\bar{x}-M(\xi)}{\frac{s^*}{\sqrt{n}}}$ is Student-distributed (or T-distributed), and $\lambda_{st} = \frac{a\sqrt{n}}{s^*}$:

$$P(-\lambda_{st} \leq \eta_{st} \leq \lambda_{st}) = p$$

λ_{st} can be determined from the values $1 - p$ and $n - 1$, and then a can be determined from λ_{st}.

- Calculate the average: **G11: =AVERAGE(C12:C31)**
- Calculate the sample standard deviation: **G12: =STDEV.S(C12:C31)**
- Calculate the number of measurements : **G13:=COUNT(C12:C31)**
- The significance level: **G14:=0.95**
- Find λ_{st} for p and n: **G16:= T.INV.2T(1-G14;G13-1)**
- The confidence limit: $a = \lambda_{st} * \frac{s^*}{\sqrt{n}}$
  - **G18:= G16*G12/SQRT(G13)**
- Then the answer is: The filling weight of a detergent is 1.52 ± 0.043 kg with the confidence level 95 %.

By increasing n the confidence limit can be decreased.

- Calculate the 2% relative error (this the 2% of the average, h(2%)): **G21: =G11*0,02**

- To find n for which a is lower than the relative error make a new table with the following headings:
  - **F23:=n**
  - **G23:=lambda_student**
  - **H23:=a**
- First, in the first column write for example 5, 10 and 15: **F24:=5; F25:=10; F26:=15**
- For each number of measurements calculate λ_{st} and a:
  - **G24: =** T.INV.2T **(1-G$14;(F24-1))**
  - **H24:= =G24*G$12/SQRT(F24)**
  - And pull them down.
- Fill out the table with increasing n and calculate a up until it gets lower than h(2%).
- The value of a gets lower than h(2%).
- (n=38 measurements are required to reduce the relative error below 2%.)

3. Exercise

A fluid with a density of $\rho = 1000 \pm 15\ kg/m^3$ is filled into a cylindrical container. The diameter of the cylinder is $d = 200 \pm 0.5\ mm$, the height $h = 200 \pm 1\ mm$. Calculate the absolute error of the fluid mass! (The absolute error is two times the standard deviation).

Solution:

In this exercise the desired quantity is the fluid mass, which can be derived from the height h and diameter d of the cylinder and form the density ρ of the fluid:

$$m = \rho V = \rho \frac{d^2 \pi}{4} h.$$

Firs let's fill out the excel table:

- The density of the fluid:
  - **D8:=1000 [kg/m^2]**
  - **D9:=15 [kg/m^2]**
  - **D10:=D9/2 [kg/m^2]**
- The diameter of the cylinder:
  - **D11:=200/1000 [m]**
  - **D12:=0.5/1000 [m]**
  - **D13:=D12/2 [m]**
- The height of the cylinder:
  - **D14:=200/1000 [m]**
  - **D15:=1/1000 [m]**
  - **D16:=D15/2 [m]**
- Then the fluid mass without the random errors (m): **D18:=D8*D11^2*PI()*D14/4**

In the lecture the propagation of the random errors was derived, here we only summaries the results. The relative variance of the error of $y = f(x_1, x_2, \dots, x_n)$:

$$\frac{\sigma_y^2}{y^2} = \sum_{i=1}^{n} \left(\frac{\partial f(x_1, x_2, \dots, x_n)}{\partial x_i}\right)^2 \frac{\sigma_i^2}{y^2},$$

Where σ_i^2 is the standard deviation of the variable x_i. In our exercise $x_1 = \rho, x_2 = d, x_3 = h$ and $y = m$. The derivatives are the following:

$$\frac{\partial m}{\partial d} = \rho \frac{2d\pi}{4} h = \frac{2m}{d}$$

$$\frac{\partial m}{\partial h} = \rho \frac{d^2\pi}{4} = \frac{m}{h}$$

$$\frac{\partial m}{\partial \rho} = \frac{d^2\pi}{4} h = \frac{m}{\rho}$$

Hence the absolute variance:

$$\sigma_m = m \sqrt{\left(2\frac{\sigma_d}{d}\right)^2 + \left(\frac{\sigma_h}{h}\right)^2 + \left(\frac{\sigma_\rho}{\rho}\right)^2}$$

Let's calculate the above variance:

- **H22:= D18*SQRT((2*D13/D11)^2+(D16/D14)^2+(D10/D8)^2)**
- **H23:=H22*2**

We can conclude that the mass of the fluid is: $m = 6.28 \pm 0.1\ kg$.

4. Exercise

The current flowing through a diode can be calculated as follows:

$$I = e^{\frac{qU}{kT}} - 1$$

The temperature of the diode is $T = 310$K, with a systematic error of 5K. The cut in voltage is $U = 0.4$ Volt with a systematic error of 0.005 Volt. "k" is the Boltzmann constant, q is a given constant. Calculate the relative error of the current! Which systematic error has greater effect in the resulting error?

Solution:

Fill out the excel table with respect of the given data and calculate I when there is no error at all:

- **D13:=0,4**
- **D14:=0,005**
- **D15:=310**
- **D16:=5**
- **I:D18:=EXP(H14*D13/D15)-1**=4801870 A

In the lecture the propagation of the systematic error was derived. The relative systematic error can be determined with the following formula:

$$\frac{\Delta I}{I} = \sum_{i=1}^{2} \frac{\partial I}{\partial x_i} \frac{\Delta x_i}{I}$$

Here $x_1 = U$ and $x_2 = T$. The derivatives:

$$\frac{\partial I}{\partial U} = \frac{q}{kT} e^{\frac{qU}{kT}}$$

$$\frac{\partial I}{\partial T} = -\frac{q}{kT^2} e^{\frac{qU}{kT}}$$

The relative systematic error of U:

- The numerator: **I18:=H14/D15*EXP(H14*D13/D15)*D14**=923430,7 V
- Then the relative error: **I22=G18/D18**=19%

The relative systematic error of T:

- The numerator: **K18= -H14*D13/D15^2*EXP(H14*D13/D15)*D16**=-1191524
- Then the relative error: **K22=K18/D18**=-25%

And the total relative error: **G22=I22+K22**=-5,58%