STATISTISCHE METHODEN DER DATENANALYSE

Excersise Sheets

Physik

UNI WIEN

10. November 2023

1 Sheet

1.1 Introduction

1.1.1 Minimum Value of $P(A \cap B)$

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) \tag{1.1}$$
$$\Leftrightarrow P(A \cap B) = P(A) + P(B) - P(A \cup B) \tag{1.2}$$

To minimize $P(A \cup B)$ maximize $P(A \cup B)$! Since $P(A) + P(B) \geq 1$ we can assume that $A \cup B = \Omega$ in effort to maximize $P(A \cup B)$. Therefore set $P(A \cup B) = 1$. Now

$$\min_{A,B} P(A \cap B) = P(A) + P(B) - P(A \cup B) = \frac{1}{21} \tag{1.3}$$

1.1.2 A and B independent

For A and B independent

$$P(A \cap B) = P(A)P(B) = \frac{5}{21} \tag{1.4}$$

1.1.3 C:= None of the Events occour

The Negations of Events are independent for independent Events. So,

$$P(A' \cap B') = P(A')P(B') = \frac{2}{3}\frac{2}{7} = \frac{4}{21} \tag{1.5}$$

1.1.4 D:= Exactly one of the two events occours

Lets define $AB := A \cap B$.

$$P(A'B \cup AB') = P(A'B) + P(AB') - P\Bigg(\underbrace{A'B \cap AB'}_{=\emptyset}\Bigg) \tag{1.6}$$
$$= P(A'B) + P(AB') \tag{1.7}$$
$$= P(A')P(B) + P(A)P(B') \tag{1.8}$$
$$= \frac{2}{3}\frac{5}{7} + \frac{1}{3}\frac{2}{7} = \frac{4}{7} \tag{1.9}$$

1.1.5 E:= Both Events occour

For A and B independent

$$P(A \cap B) = P(A)P(B) = \frac{1}{3}\frac{5}{7} = \frac{5}{21} \tag{1.10}$$

Sanity Check Since the already calculated Events $C :=$ **No Event occuring**, $D :=$ **One Event occuring** and $E :=$ **Both events occuring** span Ω, we can check that

$$P(C) + P(D) + P(E) = \frac{4}{21} + \frac{4}{7} + \frac{5}{21} \stackrel{!}{=} 1 \tag{1.11}$$

and indeed it is.

1.1.6 F:= At least one of the two Events occours

Fast Route Since $F = D \dot\cup E$

$$P(F) = P(D) + P(E) = \frac{17}{21} \tag{1.12}$$

Long, but instructive Route Again, using $AB := A \cap B$,

$$\begin{aligned}
P(D \cup E) &= P\left(A'B \cup AB' \cup AB\right) && (1.13)\\
&= P\left(A'B\right) + P\left(AB' \cup AB\right) - P\left(\underbrace{A'B \cap \left(AB' \cup AB\right)}_{=\emptyset}\right) && (1.14)\\
&= P\left(A'B\right) + P\left(AB'\right) + P(AB) - P\left(\underbrace{AB' \cap AB}_{=\emptyset}\right) && (1.15)\\
&= P\left(A'\right) P(B) + P(A) P\left(B'\right) + P(A) P(B) && (1.16)\\
&= \frac{2}{3}\frac{5}{7} + \frac{1}{3}\frac{2}{7} + \frac{1}{3}\frac{5}{7} && (1.17)\\
&= \frac{17}{21} && (1.18)
\end{aligned}$$

1.1.7 G:= At most one of the two Events occours

Fast Route Since $G := C \dot\cup D$

$$P(G) = P(C) + P(D) = \frac{16}{21} \tag{1.19}$$

Long (shortened), but instructive Route Using the same empty Intsection argument as in 1.6, 1.14 and 1.15

$$\begin{aligned}
P\left(A'B' \cup A'B \cup AB'\right) &= P\left(A'B'\right) + P\left(A'B\right) + P\left(AB'\right) && (1.20)\\
&= \frac{2}{3}\frac{2}{7} + \frac{2}{3}\frac{5}{7} + \frac{1}{3}\frac{2}{7} = \frac{16}{21} && (1.21)
\end{aligned}$$

1.2 Novel Detector Failure

There are two remoTES, each fully functional (c) with a probability

$$P(c) = P(p_1)\,P(w)\,P(p_2) = 0.405 \tag{1.22}$$

, since the sucess events of three components (p_1, w and p_2) are independent. Each one will fail with a probability

$$P(\overline{c}) = 1 - 0.405 = 0.595 \tag{1.23}$$

The probability for both to fail is

$$P(\overline{c_1} \cap \overline{c_2}) = P^2(\overline{c}) = 0.354025 \tag{1.24}$$

Which means at least one will function with probability

$$P(c_1 \cup c_2) = 1 - 0.354025 = 0.645975 \tag{1.25}$$

1.3 graduation rate

1.3.1 graduation rate

With the Events $w :=$ **is Woman**, $g :=$ **graduated** the probability of degree completion is

$$\begin{aligned} P(g) &= P(g|w)\,P(w) + P(g|\overline{w})\,P(\overline{w}) && (1.26)\\ &= 0.372 \cdot 0.625 + 0.311 \cdot (1 - 0.625) = 0.349125 && (1.27) \end{aligned}$$

1.3.2 percentage of male dropouts

Since we now know the unconditional dropout rate, we can calculate

$$\begin{aligned} P(m|\overline{g}) &= \frac{P(m \cap \overline{g})}{P(\overline{g})} && (1.28)\\ &= \frac{P(\overline{g}|m)\,P(m)}{P(\overline{g})} && (1.29)\\ &= \frac{(1 - 0.311)\,(1 - 0.625)}{1 - 0.349125} && (1.30)\\ &\approx 0.397 && (1.31) \end{aligned}$$

2 Sheet

2.1 Bad Detector

It makes sense to assume, that if an event is only registered with probability p, this can be translated to a reduced rate $\lambda_r = p\lambda$. Therefore the Number of registered Events is poisson distributed with density

$$f(k;\lambda_r) = \frac{\lambda^k}{k!}e^{-\lambda} \tag{2.1}$$

yes, but rigorous?

2.2 Uneconomical Warranty for faulty computer monitors

2.2.1 Average faultless running time for economical warranty

The percentage of failed monitors after t years is

$$\begin{aligned} p_f &= \int_0^t f_{\mathrm{EX}}\left(t';\tau\right)\, dt' && (2.2)\\ &= \int_0^t \frac{1}{\tau}e^{-\frac{t}{\tau}}\, dt' && (2.3)\\ &= 1 - e^{-\frac{t}{\tau}} && (2.4)\end{aligned}$$

For $t = 5$ and $p_f = 0.2$:

$$\begin{aligned} &0.2 \geq 1 - e^{-\frac{5}{\tau}} && (2.5)\\ \Leftrightarrow &0.8 \leq e^{-\frac{5}{\tau}} && (2.6)\\ \Leftrightarrow &-\frac{5}{\ln(0.8)} \leq \tau && (2.7)\\ \Leftrightarrow &22.4 \leq \tau && (2.8)\end{aligned}$$

2.2.2 Shortened Warranty

$$\tau \geq -\frac{3}{\ln(0.8)} = 13.44 \tag{2.9}$$

2.2.3 Monitors running after 9 years

Assuming at least 90% run after 3 years:

$$\tau \geq -\frac{3}{\ln(0.9)} = 28.47 \tag{2.10}$$

The percentage failed monitors after 5 years is

$$p_f = 1 - e^{-\frac{t}{\tau}} \tag{2.11}$$
$$= 0.161 \tag{2.12}$$

After 5 years 0.83% are still running.

2.3 Vulcano eruption

Let the waiting times to eruption for the N identical vulcanos be realized within the random variables $X_1, X_2, \ldots, X_N$. Each one distributed exponentially

2.3.1 Mean Time to eruption

2.3.2 Time until first eruption

For each X_i:

$$P(X_i \leq t) = \int_0^t \frac{1}{\tau} e^{-\frac{t'}{\tau}}\, dt' = 1 - e^{-\frac{t}{\tau}} \tag{2.13}$$
$$\Rightarrow P(X_i > t) = e^{-\frac{t}{\tau}} \tag{2.14}$$

The time until the first eruption Y is now distributed as

$$P(Y \leq t) = P(\min\{X_1, X_2, \ldots, X_N\} \leq t) \tag{2.15}$$
$$= 1 - P(X_1 > t, X_2 > t, \ldots, X_N > t) \tag{2.16}$$
$$= 1 - e^{-\frac{nt}{\tau}} \tag{2.17}$$

Y is distributed exponentially with mean $\frac{\tau}{n}$.

2.3.3 Time until last eruption

2.4 Electric Current

With

$$\frac{\mathrm{d}R}{\mathrm{d}I} = -\frac{U}{I^2} \tag{2.18}$$

and the Variance of R

$$\mathrm{Var}\left[R(I)\right] \approx R'\left(\mathrm{E}\left[I\right]\right)^2 |_{I,U} \cdot \mathrm{Var}\left[I\right] \tag{2.19}$$
$$\Rightarrow \Delta R = \left|\frac{\mathrm{d}R}{\mathrm{d}I}\right| \Delta I \tag{2.20}$$

the variance of the resistance in linear approximation is given by

$$\Delta R = \frac{230\text{V}}{2.5^2\text{A}^2} 0.15\text{A} = 5.52\Omega \tag{2.21}$$

The relative error ist then

$$\frac{\Delta R}{R} = 0.06 \tag{2.22}$$

2.4.1 Different Resistor

Halving the Resistance means halving the voltage or doubling the current.

$$R' = \frac{1}{2}\frac{U}{I} \tag{2.23}$$

Both lead to different outcomes. Assuming the voltage halves:

$$\Delta R = \frac{115\text{V}}{2.5^2\text{A}^2} 0.15\text{A} = 2.76\Omega \tag{2.24}$$

$$\Rightarrow \frac{\Delta R}{R} = 0.06 \tag{2.25}$$

, because U is linear in the error, as is the resistance.

3 Sheet

Inverse Distribution

Let X be gamma distributed $\text{Ga}\left(a,b\right), a > 2$ and $Y = 1/x$.

3.0.1 Density of $1/X$

Short Route Since $Y = h\left(X\right) = 1/X \Longrightarrow X = h^{-1}\left(Y\right) = \frac{1}{Y}$

$$f_Y(y) = \frac{1}{X^2} f_X\left(\frac{1}{X}\right) \tag{3.1}$$

Long Route The Distribution function of $Y = 1/X$ can be written as

$$P\left(\frac{1}{X} \leq x\right) = P\left(\frac{1}{x} \leq X\right) = 1 - P\left(X < \frac{1}{x}\right) = 1 - F_{\text{Ga}}\left(\frac{1}{x}; a, b\right) \tag{3.2}$$

From this the density can be retrieved by derivating with respect to y:

$$\frac{\partial}{\partial x}\left[1 - F_{\text{Ga}}\left(\frac{1}{x}; a, b\right)\right] = -\frac{\partial}{\partial x}\left[F_{\text{Ga}}\left(\frac{1}{x}; a, b\right)\right] = \frac{1}{x^2} f_{\text{Ga}}\left(\frac{1}{x}; a, b\right) \tag{3.3}$$

Note that no properties of the Gamma distribution have been used and this is a general result.

3.0.2 expectation value of $1/X$

The expectation value is

$$E\left[\frac{1}{X}\right] = \int_{-\infty}^{\infty} \frac{1}{x} f_{\mathrm{Ga}}\left(x;a,b\right)\,dx \tag{3.4}$$
$$= \int_{0}^{\infty} \frac{1}{x} \frac{x^{a-1}e^{-x/b}}{b^a\Gamma\left(a\right)}\,dx \tag{3.5}$$
$$= \int_{0}^{\infty} \frac{x^{a-2}e^{-x/b}}{bb^a\left(a-1\right)\Gamma\left(a-1\right)}\,dx \tag{3.6}$$
$$= \frac{1}{b\left(a-1\right)} \int_{-\infty}^{\infty} f_{\mathrm{Ga}}\left(x;a-1,b\right)\,dx \tag{3.7}$$
$$= \frac{1}{b\left(a-1\right)} \tag{3.8}$$

3.0.3 variance of $1/X$

The variance is

$$V\left[\frac{1}{X}\right] = E\left[\frac{1}{X^2}\right] - E\left[\frac{1}{X}\right]^2 \tag{3.9}$$
$$= \frac{1}{b^2\left(a-2\right)\left(a-1\right)} - \frac{1}{b^2\left(a-1\right)^2} \tag{3.10}$$
$$= \frac{\left(a-1\right)}{b^2\left(a-2\right)\left(a-1\right)^2} - \frac{\left(a-2\right)}{b^2\left(a-2\right)\left(a-1\right)^2} \tag{3.11}$$
$$= \frac{1}{b^2\left(a-2\right)\left(a-1\right)^2} \tag{3.12}$$

Linear Error propagation

$$\frac{1}{X} \approx \frac{1}{\mu} - \frac{1}{\mu^2}\left(x-\mu\right) + \frac{1}{\mu^3}\left(x-\mu\right)^2 \tag{3.13}$$

$$E\left[\frac{1}{X}\right] \approx \frac{1}{\mu} + \frac{1}{2}\frac{2}{\mu^2} \cdot V\left[X\right] \tag{3.14}$$
$$= \frac{1}{ab} + \frac{1}{a^2b^2} \cdot ab^2 \tag{3.15}$$
$$= \frac{1}{ab} + \frac{1}{a} \tag{3.16}$$

3.1 Energy spectrum

3.1.1 convolution

$$f_m = \frac{1}{10}\left[9 f_{No}\left(\mu = 5.9, \sigma_0^2\right) + f_{No}\left(\mu = 6.49, \sigma_1^2\right)\right] \star f_{No}\left(\mu = 0, \sigma_F^2\right) \tag{3.17}$$

$$= \frac{9}{10}\left[f_{No}\left(\mu = 5.9, \sigma_0^2\right) \star f_{No}\left(\mu = 0, \sigma_F^2\right)\right] \tag{3.18}$$

$$+ \frac{1}{10}\left[f_{No}\left(\mu = 6.49, \sigma_1^2\right) \star f_{No}\left(\mu = 0, \sigma_F^2\right)\right] \tag{3.19}$$

$$= \frac{9}{10} f_{No}\left(\mu = 5.9, \sigma_0^2 + \sigma_F^2\right) + \frac{1}{10} f_{No}\left(\mu = 6.49, \sigma_1^2 + \sigma_F^2\right) \tag{3.20}$$

3.1.2 natural uncertainties

$$\sigma^2 = \tilde{\sigma}^2 - \sigma_F^2 = 0.15\text{keV}^2 \Rightarrow \sigma = 0.39\text{keV} \tag{3.21}$$