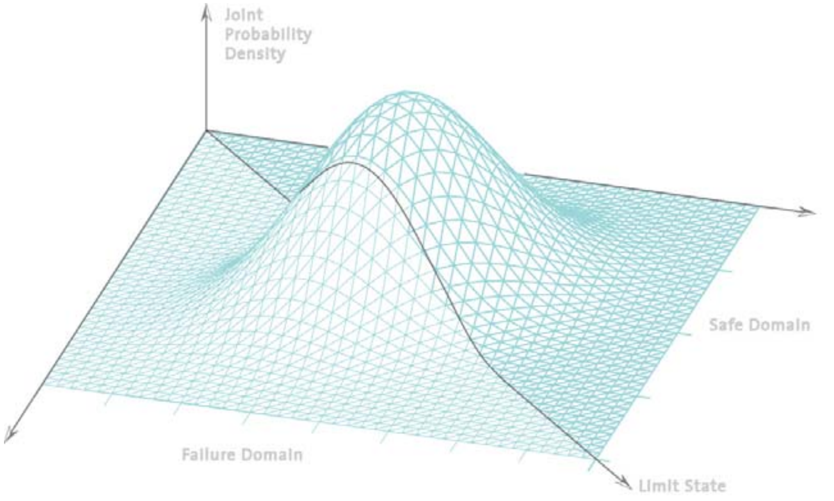


Exercises Tutorial 7

Statistics and Probability Theory

Prof. Dr. Michael Havbro Faber
Swiss Federal Institute of Technology Zurich
ETHZ

Assessment 1

- Corrections are still ongoing
- Results to be announced on the web page on the 22nd of May
- Multiple choice 1.13 (the one with the pipelines) will be omitted (many found the text “imprecise”)
- Multiple choice 1.23 (the last one) will be omitted – Printing error…there was actually no correct answer.
- So your mark will be based on 21 multiple choice questions (instead of 23)

Assessment 1

- The Assessment itself and the solution will be published in the web page at the END of the semester

- You are welcomed to come during the office hours or after making an appointment and check your own assessment – after May 22

  (identify the areas where you are strong and those where you need to pay attention)

- An overall discussion of the assessment results will follow on the next exercise Tutorial, May 24

Exercise 6.3 (Group exercise)

What is given? Success rate = 27% → $p = 0.27$

a. How large is the probability that the company will have at least one success after 12 project proposals?

What is required: Trials till first success – or number of successes?
How can you express "at least"?

b. How large is the probability that only the last of 10 project proposals is accepted?

What is required: Trials till first success – or number of successes?

c. How large is the probability that at most 2 out of 13 project proposals are accepted?

What is required: Trials till first success – or number of successes?
How can you express "at most"?

Exercise 6.3 (Group exercise)

What is given? Success rate = 27% → $p = 0.27$

a. How large is the probability that the company will have at least one success after 12 project proposals?

What is required: Trials till first success – or $\boxed{\text{number of successes}}$?
How can you express "at least"?
Binomial distribution

At least: One or more successes; anything else than 0 successes out of 12 proposals
→ complementary event

Exercise 6.3 (Group exercise)

What is given? Success rate = 27% → $p = 0.27$

b. How large is the probability that only the last of 10 project proposals is accepted?

What is required: Trials till first success – or number of successes?

Geometric distribution

Exercise 6.3 (Group exercise)

What is given? Success rate = 27% → $p = 0.27$

c. How large is the probability that at most 2 out of 13 project proposals are accepted?

What is required: Trials till first success – or $\boxed{\text{number of successes}}$?
How can you express "at most"?

Binomial distribution

At most: Or zero, or one, or two successes.

Exercise 6.2

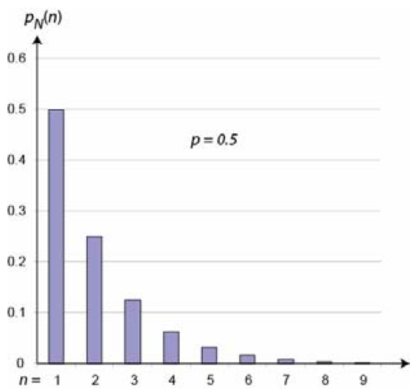
A dike is designed to withstand a “1000-year return period flood”. How large is the probability that water will overflow the dike in the following conditions:

a) During a 10 year period, for the first time in a given year?
b) During a 10 year period, twice?
c) Will not overflow during a 10 year period?
d) During a 10 year period, at most once?
e) During a 100 year period, 10 times?
f) -
g) During a 1000 year period, once or more often?

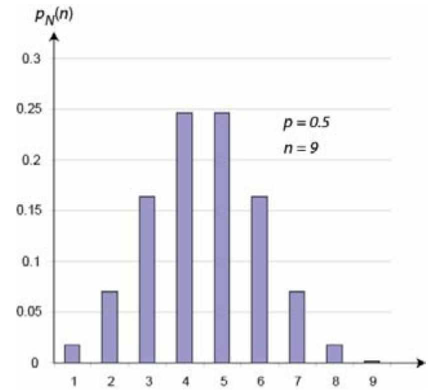
(It is assumed that flood occurs at most once in a year.)

Exercise 6.2

Geometric distribution
Time till first success



Binomial distribution
Number of success



How large is the probability that water will overflow the dike …

a) … during a 10 year period, **for the first time** in a given year (for example, in the 10^{th} year) ?
→ Geometric distribution:

$$P(H_{overflow,1}) = (p)(1-p)^{n-1}$$

$$= (0.001)(0.999)^9 = 0.000991$$

b) … during a 10 year period, **twice**? → Binomial distribution

$$P[Y = y] = \binom{n}{y} p^y (1-p)^{n-y}$$

$$P(H_{overflow,2}) = \frac{10!}{2! \cdot (10-2)!}(p)^2(1-p)^{10-2}$$

$$= \frac{10 \cdot 9}{2}(0.001)^2(0.999)^8 = 0.000045$$

Exercise 6.2

A dike is designed to withstand a "1000-year return period flood". How large is the probability that water will overflow the dike in the following conditions:

c) Will not overflow in a 10 year period?

→ Binomial distribution

$$P(H_{overflow,0}) = \frac{10!}{0! \cdot (10-0)!}(p)^0 \cdot (p-1)^{10-0} = (0.001)^0 \cdot (0.999)^{10} = 0.99004$$

d) During a 10 year period, at most once?

Probability that in 10 years you get ***or 0 or 1*** overflows → Binomial distribution

$$P(H_{\max,1}) = P(H_{overflow,0}) + P(H_{overflow.,1}) = 0.99995$$

$$P(H_{overflow,1}) = \frac{10!}{1! \cdot (10-1)!}(p)^1 (p-1)^{10-1} = 10(0.001)^1 (0.999)^9 = 0.00991$$

Exercise 6.2

A dike is designed to withstand a "1000-year return period flood". How large is the probability that water will overflow the dike in the following conditions:

e) During a 100 year period, 10 times?

$$P(H_{overflow,10}) = \frac{100!}{10!(100-10)!}(p)^{10}\cdot(p-1)^{100-10} = \frac{100!}{10!(90)!}(0.001)^{10}\cdot(0.999)^{90}$$

$$= 1.73\cdot 10^{13}\cdot 10^{-30}\cdot 0.914 = 1.6\cdot 10^{-17}$$

g) During a 1000 year period, once or more often?

Probability that in 1000 years you get **any other result than 0** overflows

→ Binomial distribution

→ the required probability is the probability of the complementary event "0 overflows"

$$P(H_{overflow,0}) = \frac{1000!}{0!(1000-0)!}(p)^{0}\cdot(p-1)^{1000-0} = (0.001)^{0}(0.999)^{1000} = 0.368$$

$$P(H_{overflow,\geq 1}) = 1-0.368 = 0.632$$

Exercise 6.1

Let $\{X_i\}_{1\le i\le 50}$ be independent, identically Normal distributed with mean value of $\mu = 1$ and standard deviation of $\sigma = 2$. Define:

$$S_n = X_1 + X_2 + \ldots + X_n \qquad \text{and} \qquad \bar{X}_n = \frac{1}{n}\left(X_1 + X_2 + \ldots + X_n\right) = \frac{S_n}{n}$$

where n=50.

a) Calculate the mean and the standard deviation of S_n and $\bar{X}_n$. ✓ done

b) Calculate $P\left(E\left[X_1\right]-1 \le X_1 \le E\left[X_1\right]+1\right)$ ✓ done

c) Calculate $P\left(E\left[S_n\right]-1 \le S_n \le E\left[S_n\right]+1\right)$

d) Calculate $P\left(E\left[\bar{X}_n\right]-1 \le \bar{X}_n \le E\left[\bar{X}_n\right]+1\right)$

Exercise 6.1

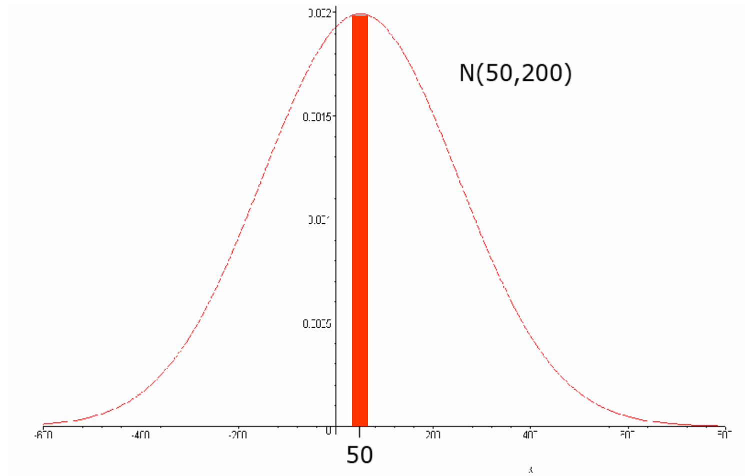
c) Calculate $P\left(E\left[S_n\right]-1 \le S_n \le E\left[S_n\right]+1\right)$

$$E\left[S_n\right]=50$$

$$P\left(E\left[S_n\right]-1 \le S_n \le E\left[S_n\right]+1\right)$$
$$= P[49 \le S_n \le 51]$$
$$= P\left[\frac{49-50}{\sqrt{200}} \le \frac{S_n-50}{\sqrt{200}} \le \frac{51-50}{\sqrt{200}}\right]$$
$$= \Phi\left(\frac{1}{\sqrt{200}}\right) - \Phi\left(-\frac{1}{\sqrt{200}}\right)$$
$$= \Phi\left(\frac{1}{\sqrt{200}}\right) - \left(1-\Phi\left(\frac{1}{\sqrt{200}}\right)\right)$$
$$= 2\cdot\Phi\left(\frac{1}{\sqrt{200}}\right) - 1 = 2\cdot\Phi(0.07)-1$$
$$= 2\cdot 0.53 - 1 = 0.06$$

$$Z = \frac{S_n-\mu}{\sigma}$$

Standardization



z	$\Phi(z)$	z	$\Phi(z)$
0.00	0.5000	0.50	0.69
0.01	0.5040	0.51	0.69
0.02	0.5080	0.52	0.69
0.03	0.5120	0.53	0.70
0.04	0.5160	0.54	0.70
0.05	0.5199	0.55	0.70
0.06	0.5239	0.56	0.71
0.07	0.5279	0.57	0.71
0.08	0.5319	0.58	0.71
0.09	0.5359	0.59	0.72
0.10	0.5398	0.60	0.72
0.11	0.5438	0.61	0.72
0.12	0.5478	0.62	0.73
0.13	0.5517	0.63	0.73

Exercise 6.1

d) Calculate $P\left(E\left[\overline{X}_n\right]-1 \le \overline{X}_n \le E\left[\overline{X}_n\right]+1\right)$

steps:
- Standardization
- Probability table

$$E\left[\overline{X}_n\right]=1$$

Calculated last time

$$V\left[\overline{X}_n\right]=0.08$$

$$P\left(E\left[\overline{X}_n\right]-1 \le \overline{X}_n \le E\left[\overline{X}_n\right]+1\right)$$

$$=P\left(0 \le \overline{X}_n \le 2\right)$$

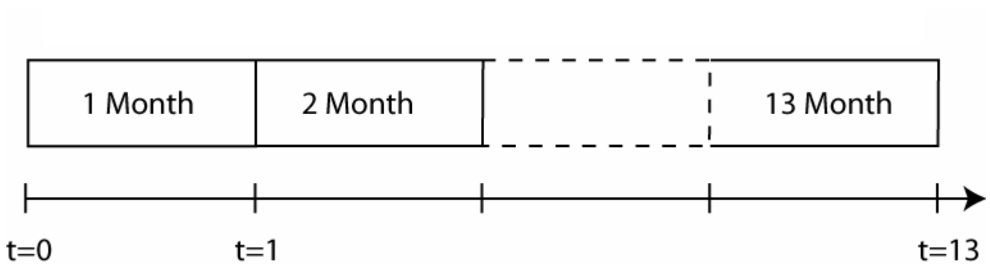
$$=P\left(\frac{0-1}{\sqrt{0.08}} \le Z \le \frac{2-1}{\sqrt{0.08}}\right)$$

$$=P\left(-3.5 \le Z \le 3.5\right)$$

z)	z	$\Phi(z)$	z	$\Phi(z)$	z	$\Phi(z)$
15	1.00	0.8413	1.50	0.9332	2.00	0.9772
50	1.01	0.8438	1.51	0.9345	2.10	0.9821356
85	1.02	0.8461	1.52	0.9357	2.20	0.9860966
19	1.03	0.8485	1.53	0.9370	2.30	0.9892759
54	1.04	0.8508	1.54	0.9382	2.40	0.9918025
88	1.05	0.8531	1.55	0.9394	2.50	0.9937903
23	1.06	0.8554	1.56	0.9406	2.60	0.9953388
57	1.07	0.8577	1.57	0.9418	2.70	0.9965330
90	1.08	0.8599	1.58	0.9429	2.80	0.9974449
24	1.09	0.8621	1.59	0.9441	2.90	0.9981342
57	1.10	0.8643	1.60	0.9452	3.00	0.9986501
91	1.11	0.8665	1.61	0.9463	3.10	0.9990324
24	1.12	0.8686	1.62	0.9474	3.20	0.9993129
57	1.13	0.8708	1.63	0.9484	3.30	0.9995166
89	1.14	0.8729	1.64	0.9495	3.40	0.9996631
22	1.15	0.8749	1.65	0.9505	3.50	0.9997674
54	1.16	0.8770	1.66	0.9515	3.60	0.9998409
86	1.17	0.8790	1.67	0.9525	3.70	0.9998922
17	1.18	0.8810	1.68	0.9535	3.80	0.9999277

Exercise 7.1

The occurrence of rainfall in an area in a year may be described by a non-homogeneous Poisson process with the intensity, namely, the mean rate of occurrence of rainfall per unit time, $\lambda(t)$, where t is defined in the interval $[0,13]$ and describes the time in a monthly ***unit*** (i.e., 4 weeks).



$$\lambda(t) = \begin{cases} \dfrac{2 \cdot t}{3} & \text{for } 0 \le t \le 3 \\ 2 & \text{for } 3 < t \le 7 \\ \dfrac{13 - t}{3} & \text{for } 7 < t \le 13 \end{cases}$$

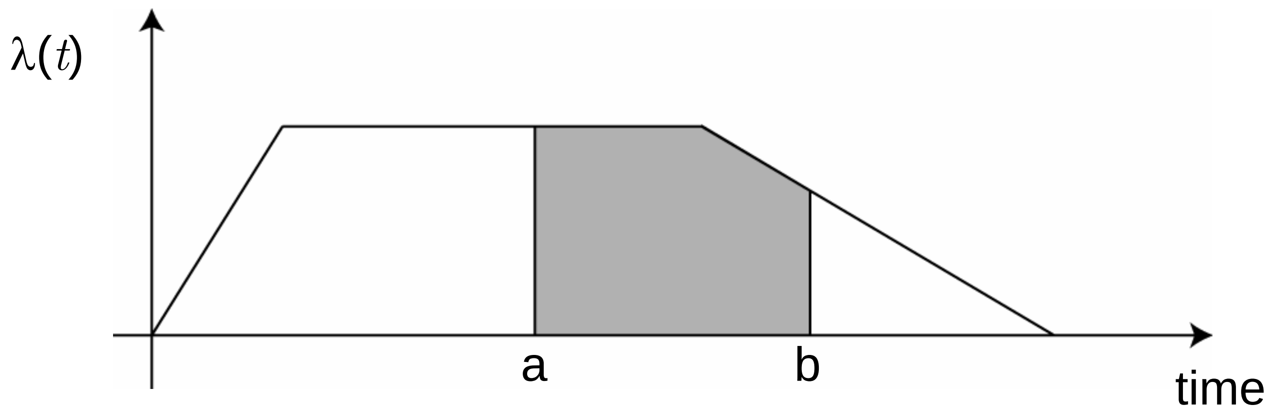
Please correct it in tutorials book

Hint: For a non-homogeneous Poisson process, the intensity varies with time.
The mean occurrence rate for any time interval (*t1, t2*) of the Poisson process can be described by:

$$\nu = \int_{t_1}^{t_2} \lambda(t)\, dt$$

Exercise 7.1

Mean occurrence rate of an event per unit time $\lambda(t)$



Mean occurrence rate in time [a, b]: ν $\qquad \nu = \int_a^b \lambda(t)dt$

$$P_n(t) = \frac{\left(\int_0^t \lambda(\tau)d\tau\right)^n}{n!}\exp\left(-\int_0^t \lambda(\tau)d\tau\right) \longrightarrow P_n(t) = \frac{\nu^n}{n!}e^{-\nu}$$

n = number of occurrence of the event
(t) = period of interest

Exercise 7.1

$$\lambda(t) = \begin{cases} \dfrac{2\cdot t}{3} & \text{for } 0 \le t \le 3 \\ 2 & \text{for } 3 < t \le 7 \\ \dfrac{13-t}{3} & \text{for } 7 < t \le 13 \end{cases}$$

a) Calculate the probability that in the first 5 months of a year, three or more rainfalls occur.

Steps:

Random variable T = number of rainfalls in the first 5 months

Obtain the parameter in the Poisson process: $\nu = \int_0^5 \lambda(t)dt = \int_0^3 \frac{2\cdot t}{3}dt + \int_3^5 2dt = \ldots$

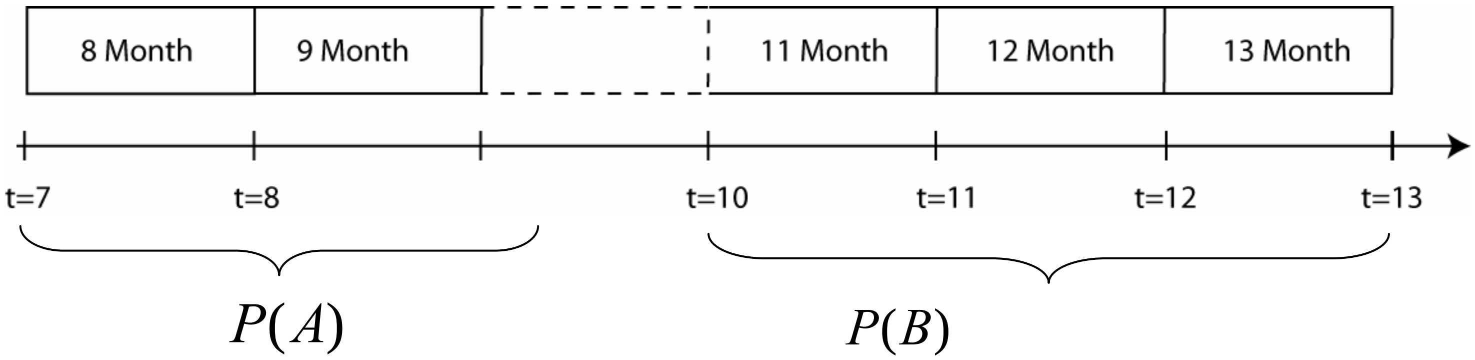
Calculate the probability:

$$\sum_{i=3}^{\infty} P_T(5) = 1 - \left[P_0(5) + P_1(5) + P_2(5)\right]$$

$$\begin{cases} P_0(5) = \dfrac{\nu^n}{n!}e^{-\nu} = \ldots \\ P_1(5) = \ldots \\ P_2(5) = \ldots \end{cases}$$

Exercise 7.1

b) Calculate the probability that a rainfall occurs
at most once during the 8th, 9th and 10th month and
at most once during the last 3 months of a year.



Let Event A represent the number of occurrences of a rainfall in the 8th, 9th, 10th month (Δt_a)

Let Event B represent the number of occurrences of a rainfall in the 11th, 12th and 13th month (Δt_b)

$$P[A \cap B] = P(A) \cdot P(B)$$ Independency!

Exercise 7.1

b) Calculate the probability that a rainfall occurs
at most once during the 8th, 9th and 10th month and
at most once during the last 3 months of a year.

$$\lambda(t) = \begin{cases} \frac{2 \cdot t}{3} & \text{for } 0 \le t \le 3 \\ 2 & \text{for } 3 < t \le 7 \\ \frac{13-t}{3} & \text{for } 7 < t \le 13 \end{cases}$$

Steps

Calculate the mean occurrence rate for each period:

$$\nu_1 = \int_7^{10} \frac{1}{3}(13-t)\,dt = \ldots$$

$$\nu_2 = \int_{10}^{13} \frac{1}{3}(13-t)\,dt = \ldots$$

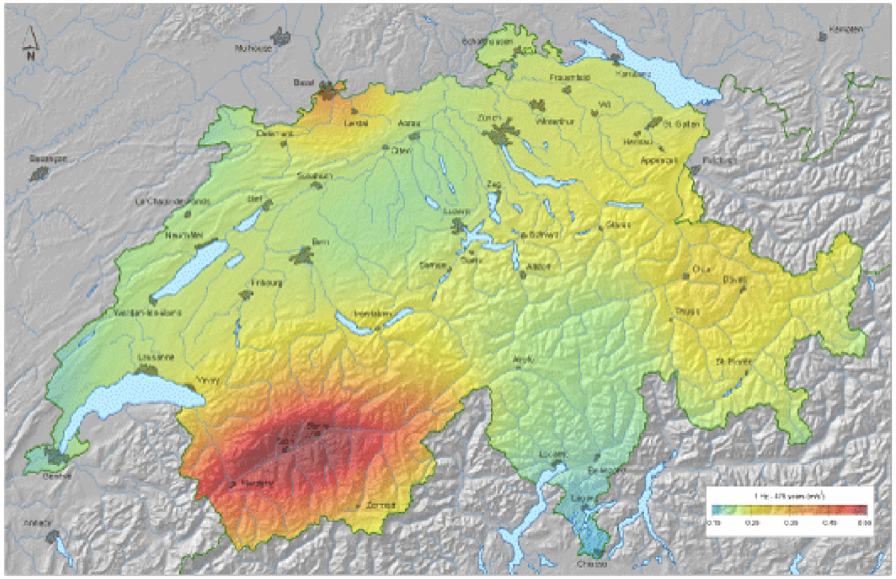
Calculate the probability of at most one rainfall (or 0, or 1) :

Exercise 7.2

An earthquake hazard map is often represented in terms of peak ground acceleration and a return period of 475 years is adopted in the map in many countries.

a. Show that the event with a return period of 475 years corresponds to the event whose occurrence probability is 10% in 50 years, under the assumption that an event follows a homogeneous Poisson process.

b. What is the probability that an earthquake with a return period of 475 years will occur within the next 475 years?



Seismic Shaking Hazards in Switzerland

Probabilistic Seismic Hazards Assessment 10% probability of being exceeded in 50 years

www.earthquake.ethz.ch

Exercise 7.2

What is given?

Mapping of earthquakes with return periods of 475 years
Earthquake event
- follows a homogeneous Poisson process
- occurrence probability is 10% in 50 years

The 10% probability of exceeding (a certain ground motion) in 50 years maps depict an annual probability of 1 in 475 of being exceeded each year (→ 90% chance that these ground motions will NOT be exceeded).

a. Verify this relationship!
Event with a return period of 475 years = Event with probability 10% in 50 years

b. What is the probability that an earthquake with a return period of 475 years will occur within the next 475 years?

Exercise 7.2

a. Verify this relationship!
Event with a return period of 475 years = Event with probability 10% in 50 years

Annual probability of an event with a return period of 475 years:

$$P_A(1) = \frac{1}{475}$$

Probability of occurring in the next 50 years is:
(probability of occurring in the 1st year + probability of occurring in the 2nd year + … + … + probability of occurring in the 50th year)

$$\sum_{i=1}^{n} p(1-p)^{n-1} = 1-(1-p)^n$$

- Geometric distribution, summed up over 50 years
- Cumulative distribution function of the Geometric distribution!

$$P_A(50) = 1-(1-P_A(1))^{50} = 1-\left(1-\frac{1}{475}\right)^{50} = 0.1$$

Exercise 7.2

b. What is the probability that an earthquake with a return period of 475 years will occur within the next 475 years?

Probability of occurring in the next 475 years is:

Type of distribution:

$P_A(475) =$

Exercise 7.2 (Think in another way…!)

What is given? Return period

So…annual probability of occurrence: (check script after Equation D.60)

Average time till "success"- (here till occurrence) (check script Equation D.59)

Time till and between events, described by the Poisson process, is Exponential distributed (script Equation D.64)

a. $P[T \leq 50\,Jahren] =$ (script Table D.1)

b. $P[T \leq 475\,Jahren] =$

Exercise 7.4 (Group exercise)

The operational time of a diesel engine until a breakdown, is assumed to follow an Exponential distribution with mean $\mu_T = 24$ months.

Normally such an engine is inspected every 6 months and in case that a default is observed this is fully repaired. It is assumed that a default is a serious damage that leads to breakdown if the engine is not repaired.

a. Calculate the probability that such an engine will need repair before the first inspection.

We are looking for: $P(T \leq 6 \text{ months}) = \ldots$

We know that the time until breakdown is exponentially distributed: $F_T(t) = 1 - e^{-\lambda t}$

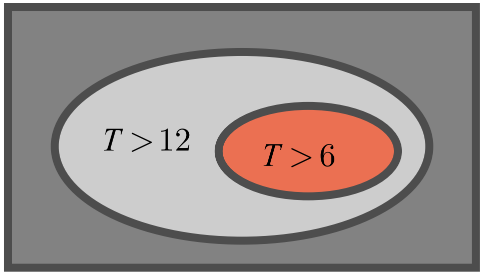
(Script page D-17)

Exercise 7.4 (Group exercise)

b. Assume that the first inspection has been carried out and no repair was required. Calculate the probability that the diesel engine will operate normally until the next scheduled inspection.

$$P(\text{no repair up to the second inspection} \mid \text{no repair at the first inspection})$$

Hint: $$P[T>12 \mid T>6] = \frac{P[T>12 \cap T>6]}{P[T>6]} = \frac{P[T>12]}{P[T>6]}$$



$0 \le P(T>6) \le P(T>12)$

Please correct in exercise tutorials book: The hint is for part b (not c)

c. Calculate the probability that the diesel engine will fail between the first and the second inspection.

$$P[6 \le T \le 12]$$

Exercise 7.4 (Group exercise)

d. A nuclear power plant owns **6 such engines**. The operational lives $T_1, T_2 \ldots T_6$ of the diesel engines are assumed statistically independent. What is the probability that at most 1 engine will need repair at the first scheduled inspection?

We are looking for: **At most** 1 repair out of 6 engines → Binomial distribution

The probability of 1 engine needing repair at the first inspection, has been calculated in a)

e. It is a requirement that the probability of repair at each scheduled inspection is not more than 60%. The operational lives $T_1, T_2 \ldots T_6$ of the diesel engines are assumed statistically independent. What should be the inspection interval?

$$P(\text{no engine needs repair at the time } t \text{ of the inspection}) = 0.6$$