Statistics II: Exercise Session 1

1 Probability Distributions

1.1 Discrete Probability Distributions

1.1.1 Hypergeometric Distribution

$X \sim H(N; M; n)$ $\quad$ $\mathrm{E}[X] = n \cdot \frac{M}{N}$ $\quad$ $\mathrm{Var}(X) = n \cdot \frac{M}{N} \cdot (1 - \frac{M}{N}) \cdot (\frac{N-n}{N-1})$

$$f_H(x; N,M,n) = \begin{cases} \dfrac{\binom{M}{x} \cdot \binom{N-M}{n-x}}{\binom{N}{n}} & \text{for } x = 0, 1, \ldots, \min(n,M); \\ 0 & \text{otherwise} \end{cases}$$

Models number of successes in n draws, *without* replacement, from a finite population of size N that contains exactly M successes.

1.1.2 Binomial Distribution

$X \sim B(n; p)$ $\quad$ $\mathrm{E}[X] = n \cdot p$ $\quad$ $\mathrm{Var}(X) = n \cdot p \cdot (1 - p)$

$$f_B(x; n,p) = \begin{cases} \binom{n}{x} \cdot p^x \cdot (1-p)^{n-x} & \text{for } x = 0, 1, \ldots, n; \\ 0 & \text{otherwise} \end{cases}$$

$$F_B(x; n,p) = \begin{cases} \sum_{k=0}^{x} \binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} & \text{for } x \geq 0; \\ 0 & \text{for } x < 0 \end{cases}$$

Models number of successes in n draws, *with* replacement, p is probability of a success in each draw.

1.1.3 Poisson Distribution

$X \sim PO(\lambda)$ $\quad$ $\mathrm{E}[X] = \lambda$ $\quad$ $\mathrm{Var}(X) = \lambda$

$$f_{PO}(x; \lambda) = \frac{\lambda^x}{x!} \cdot e^{-\lambda} \quad \text{for } x = 0, 1, 2, \ldots; \lambda > 0;$$

$$F_{PO}(x; \lambda) = \begin{cases} \sum_{k=0}^{x} \dfrac{\lambda^k}{k!} \cdot e^{-\lambda} & \text{for } x \geq 0; \lambda > 0; \\ 0 & \text{for } x < 0 \end{cases}$$

Models number of events occurring in a fixed interval of time and/or space.

1.2 Continuous Probability Distributions

1.2.1 Exponential Distribution

$X \sim EX(\lambda)$ $\quad$ $\mathrm{E}[X] = \frac{1}{\lambda}$ $\quad$ $\mathrm{Var}(X) = \frac{1}{\lambda^2}$

$$f_{EX}(x;\lambda) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{for } x \geq 0, \lambda > 0; \\ 0 & \text{for } x < 0 \end{cases}$$

$$F_{EX}(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & \text{for } x \geq 0, \lambda > 0; \\ 0 & \text{for } x < 0 \end{cases}$$

Models the time between events in a process in which events occur continuously and independently at a constant average rate (intensity).

1.2.2 Normal Distribution

$X \sim N(\mu;\sigma)$ $\quad$ $\mathrm{E}[X] = \mu$ $\quad$ $\mathrm{Var}(X) = \sigma^2$

$$f_N(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right) \quad \text{for } -\infty < x < +\infty, \sigma > 0$$

$$F_N(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp\left(-\frac{(t-\mu)^2}{2\sigma^2}\right) dt$$

1.2.3 Standard Normal Distribution

$Z \sim N(0;1)$ $\quad$ $\mathrm{E}[Z] = 0$ $\quad$ $\mathrm{Var}(Z) = 1$

$$\varphi(z) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{z^2}{2}\right) \quad \text{for } -\infty < z < +\infty$$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{z} \exp\left(-\frac{v^2}{2}\right) dv$$

Relationship between normal and standard normal distribution

$$X = \mu + Z \cdot \sigma \quad \Leftrightarrow \quad Z = \frac{X-\mu}{\sigma}$$

$$\begin{aligned} \mathrm{P}(X \leq x) \quad &= \mathrm{P}\left(\frac{X-\mu}{\sigma} \leq \frac{x-\mu}{\sigma}\right) = \mathrm{P}\left(Z \leq \frac{x-\mu}{\sigma}\right) \\ &= \Phi\left(\frac{x-\mu}{\sigma}\right) = \Phi(z) = \mathrm{P}(Z \leq z) \end{aligned}$$

Exercises

5-10: *Eggs*

About a certain type of eggs it is known that in a package of 6 eggs there are always 2 spoiled. To check their freshness, we randomly pick 3 of them and throw them in a pan.

a) What is the probability that exactly 1 out of those 3 is spoiled?

b) What is the probability that at most 1 is spoiled?

c) What is the probability that exactly 3 are spoiled?

d) How many spoiled eggs are to be expected in the sample of the three eggs?

At a one small chicken farm there were more than 500 eggs produced over a certain amount of time. It is known that with a probability of 80 % such an egg is not spoiled. There was an order of 20 eggs, which were picked randomly and delivered.

e) What is approximately the probability that more than 2 eggs in the delivery are spoiled?

f) Compute the expected value of a random variable "Number of not spoiled eggs in the delivery".

g) Calculate an approximate probability that exactly 16 eggs in the delivery are spoiled.

5-11: *Phone calls*

The average number of phone calls started by a call center between 2 and 4 p.m. is exaclty 2.5 per minute.

a) What is the distribution of a random variable X: "Number of started phone calls per minute" during the given time?

b) What is the probability that in one particular minute during this time there are

   - no;
   - less than 3;
   - 4 or more

   phone calls started?

5-14: *Electronic component*

For an electronic component one can expect 48 failures a day (= 24 hours). The failures are short-lasting and occur randomly and independently from each other.

a) What is the distribution of a time passed (in hours) between two failures? State the distribution type and its parameters.

b) What is the probability that more than 2 hours will pass until the next failure occurs?

c) Using given example, verbalize the following form:

$$\int_1^2 2e^{-2x}\,dx$$

d) Assume that an electronic system consists of two of such components which work independently from each other. The system collapses as soon as one of the components stops working. What is the probability that the system works for more than 2 hours?

5-16: *Steel pins*

A machine produces steel pins. Unfortunately, diameter of the produced pins fluctuates. The random variable X_1: "Diameter of a pin" is normally distributed with $\mu_1 = 6$ mm and $\sigma_1 = 0.4$ mm.

a) What is the probability that the diameter of a pin differs from the size of 6 mm by more than 2 %?

b) What is the probability that the diameter of a pin is exactly 6 mm?

c) Which size will not be exceeded with the probability of 85 %?

A second machine, which works independently from the first one, makes holes in a workpiece where the steel pins should be used. The diameter of the holes fluctuates, too. The random variable X_2: "Diameter of a hole" is normally distributed with $\mu_2 = 6.05$ mm and $\sigma_2 = 0.3$ mm.

d) What is the probability that the diameter of a hole is less than 6 mm?

e) What is the probability distribution of a r.v. $Y = X_2 - X_1$? (Distribution type with its parameters.)

f) What is the probability that the pin doesn't fit in the hole?