Statistics - Exam preparation

xnacly - July 30, 2023- source

Contents

1 Introduction

Statistics is a fairly big field. Therefore this paper will only include the absolutely necessary topics for passing the university class.

2 Abstract

This paper starts of with symbols used in the field of statistics, their meaning and in what context they are commonly used. Following Combinatorics is thematized.

3 Symbols and special characters

- $n!$ Factorial / Fakultät
- $\binom{n}{k}$ Binomial Coefficient / Binomialkoeffizient
- Ω Event set / Ergebnismenge
- ω Result / Ergebnis
- $A \subseteq \Omega$ Event / Ereignis
- $\{\omega\}$ Elementary event / Elementarereignis
- $\mathbb{P}$ Probability measure / Wahrscheinlichkeitsmaß
- $\mathbb{P}(A)$ Event propability / Wahrscheinlichkeit eines Ereignisses
- $\mathbb{E}(X), \mu_x, \mu$ Expected value / Erwartungswert
- σ Standard deviation / Standardabweichung
- $\mathrm{Var}(X), \sigma_x^2$ Variance / Varianz
- $\mathrm{Cov}(X, Y), \sigma_{XY}$ Kovarianz von X und Y
- $\mathcal{N}(\mu, \sigma^2)$ Normal distribution / Normalverteilung
- φ Bell curve / Glockenkurve
- Φ Error function / Fehlerintegral
- X Random variable / Zufallsvariable
- Z Standard score / standard-normalverteilte Zufallsvariable [1]

[1] read more: wikipedia

- $\text{Bin}(n,p)$ Binomial distribution / Binomialverteilung
- $\text{Pois}(\lambda)$ Poisson distribution / Poission-Verteilung
- $\text{Exp}(\lambda)$ Exponential distribution / Exponentialverteilung

4 Combinatorics

This chapter will introduce the Binomial Coefficient, Factorials, Pascal's triangle and the binomial theorem.

4.1 Binomial Coefficient

n choose k; used to calculate the Amount of Sets in $\{1, ..., n\}$ with exactly k Elements. n needs to be positive and k and n have to meet the following criteria: $n \in \mathbb{N}, 0 \leq k \leq n$.

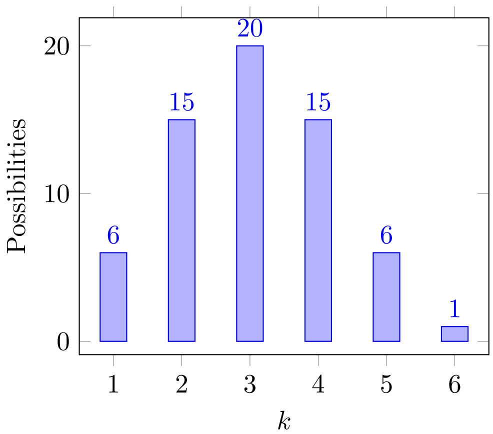
$$\binom{n}{k} = \frac{n!}{k! \cdot (n-k)!} \tag{1}$$

Choosing k different Numbers from $\{1, ..., n\}$, there are n possibilities for the first number, $n-1$ possibilities for the second, and so forth.

4.1.1 Example:

If we were to calculate how many possibilities there are to choose 3 out of 6 dishes, we can simply use the above formular:

$$\binom{6}{3} = \frac{6!}{3! \cdot (6-3)!} = \frac{720}{3! \cdot 3!} = \frac{720}{36} = \underline{\underline{20}}$$



4.2 Factorial

This can be described with $n!$. The factorial is defined as the product of decrementing n by an increasing subtrahend:

$$n! := n \cdot (n-1) \cdot (n-2)... \tag{2}$$

If $n = k$, there are $n!$ possibilities to choose k Elements from n. $0! = 1$.

4.3 Pascal's triangle

The pascal's triangle can be used to visualize the binomial coefficient.[2]

$$\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1} \tag{3}$$

4.4 Binomial Theorem

Allows for expressing the exponents of $(x+y)^n, n \in \mathbb{N}$ as a polynomial with the degree of n.

$$(x+y)^n = \sum_{k=0}^{n} \binom{n}{k} \cdot x^k \cdot y^{n-k} \tag{4}$$

5 Probability theory

This chapter contains information on how to calculate probabilities.

5.1 Event set

The set containing results of the *experiment* E is notated via the *event set* (Ω). Sub sets of Ω are *events* (ω). *Events* with one entry are *elementary events* $\{\omega\}$. If Ω is finite: $\forall \omega \in \Omega, \mathbb{P}(\omega) \geq 0$. The sum of all probabilities of $\omega \in \Omega$ is 1. [3]

5.2 Random variable

$X : \Omega \to \mathbb{R}$ we define $\{X = x\} := \{\omega | X(\omega) = x\}$ and can therefore shorten our definition of the probability that X is x to: $P(X = x) := P(\{X = x\})$

$$x \to \mathbb{P}(X = x) \tag{5}$$
$$x \to \mathbb{P}(X \leq x) \tag{6}$$

[2] read more about *pascal's triangle* here: wikipedia

[3] $\sum_{\omega \in \Omega} \mathbb{P}(\omega) = 1$

The first equation defines the density / probability function of X and the second equation the distribution function of X.

5.3 Expected value

$$\mathbb{E}(X) = \sum_{k \in \mathbb{R}} k \cdot \mathbb{P}(X = k) \tag{7}$$

5.3.1 Example:

For a dice:

$$\begin{aligned}
\mathbb{P}(X = k) &= \frac{1}{6}; k = 1, 2, 3, 4, 5, 6 \\
\mathbb{E}(X) &= \sum_{k \in 1}^{6} k \cdot \mathbb{P}(X = k) \\
&= \sum_{k=1}^{6} k \cdot \frac{1}{6} \\
&= \frac{1}{6}(1 + 2 + 3 + 4 + 5 + 6) \\
&= \left\lfloor \frac{21}{6} \right\rfloor = \left\lfloor \frac{7}{2} \right\rfloor = \underline{\underline{3}}
\end{aligned}$$

As shown above the expected value is $\lfloor \frac{7}{2} \rfloor \approx 3$. [4]

5.4 Variance

X: Random variable, $\mu = \mathbb{E}(X)$: expected value.

$$\mathrm{Var}(X) = \mathbb{E}\left\lfloor (X - \mu)^2 \right\rfloor \tag{8}$$

5.5 Standard deviation

$$\sigma_X := \mathrm{SD}^5(X) := +\sqrt{\mathrm{Var}(X)} \tag{9}$$

5.5.1 Standardized Variable

X is standardized, if $\mathbb{E}(X) = 0, \sigma_X^2 = 1$.

6 Binomial distribution

A random variable X is binomial distributed with $n \in \mathbb{N}$ and $p \in [0, 1]$ if

$$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \tag{10}$$

[4] $\lfloor a \rfloor$ denotes the flooring of a.
[5] short for standard deviation

We can now denote:

$$\begin{aligned}
X &\sim \mathrm{Bin}(n, p) && (11) \\
\mathbb{E}(X) &= np && (12) \\
\mathrm{Var}(X) &= np(1 - p) && (13)
\end{aligned}$$

Binomial distribution can be assumed if an Event either occurs or it doesn't.

6.1 Independent random binomial distributed variables

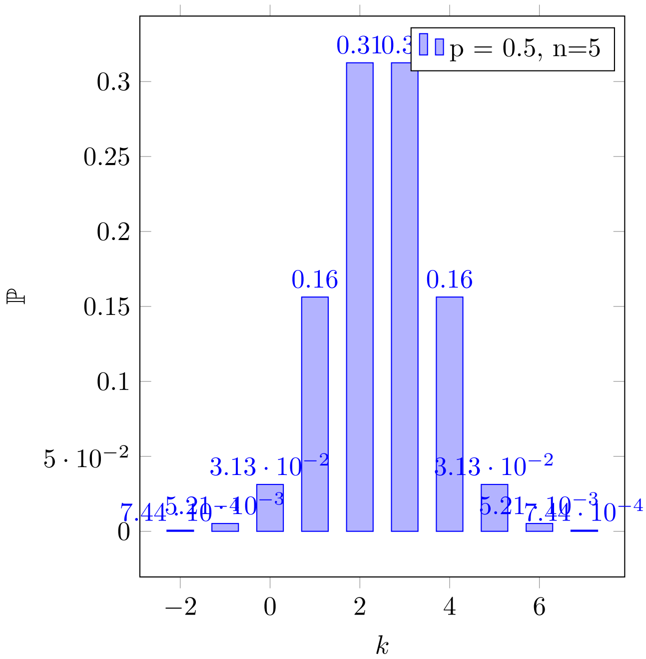
$X \sim \mathrm{Bin}(n, p)$ and $Y \sim \mathrm{Bin}(m, p)$ can be merged:

$$X + Y \sim \mathrm{Bin}(n + m, p) \tag{14}$$

6.2 Example

Calculating how probable it is to throw a penny five times and have them result in four heads.

$$\begin{aligned}
p &= 0,5 \\
\mathbb{P}(X = k) &= \binom{n}{k} p^k (1 - p)^{n-k} \\
\mathbb{P}(X = 4) &= \binom{5}{4} 0,5^4 (1 - 0,5)^{5-4} \\
&= 5 \cdot 0,0625 \cdot 0,5 \\
&= \underline{\underline{0,15625}} \to 15,625\%
\end{aligned}$$



7 Geometric distribution

A random variable $X \in \mathbb{N}$ is geometric distributed with $p \in]0,1]$ if density is

$$\mathbb{P}(X = k) = p(1-p)^{k-1} \tag{15}$$

or distributed with

$$\mathbb{P}(X \leq k) = 1 - (1-p)^k \tag{16}$$

We can denote:

$$X \sim \mathrm{G}(p) \tag{17}$$

$$\mathbb{E}(X) = \frac{1}{p} \tag{18}$$

$$\mathrm{Var}(X) = \frac{1}{p}\left(\frac{1}{p} - 1\right) \tag{19}$$

8 Poisson distribution

A random variable X is poisson distributed with $\lambda > 0$, if

$$\mathbb{P}(X = k) = \frac{\lambda^k}{k!}e^{-\lambda} \tag{20}$$

We denote:

$$X \sim \mathrm{Pois}(\lambda) \tag{21}$$

$$\mathbb{E}(X) = \lambda \tag{22}$$

$$\mathrm{Var}(X) = \lambda \tag{23}$$

8.1 Independent random poisson distributed variables

$X \sim \mathrm{Pois}(\lambda)$ and $Y \sim \mathrm{Pois}(\mu)$ can be merged:

$$X + Y \sim \mathrm{Pois}(\lambda + \mu) \tag{24}$$

9 Exponential distribution

A random variable $X \in \mathbb{R}_{\geq 0}$is exponential distributed with $\lambda > 0$ if density is

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & x \leq 0 \\ 0 \end{cases} \tag{25}$$

or distributed with

$$F(x) = \begin{cases} 1 - e^{-\lambda x} & x \leq 0 \\ 0 \end{cases} \tag{26}$$

We denote:

$$X \sim \mathrm{Exp}(\lambda) \tag{27}$$

$$\mathbb{E}(X) = \frac{1}{\lambda} \tag{28}$$

$$\mathrm{Var}(X) = \frac{1}{\lambda^2} \tag{29}$$

9.1 Independent random exponential distributed variables

$X \sim \mathrm{Exp}(\lambda)$ and $Y \sim \mathrm{Exp}(\mu)$ can be merged:

$$\min(X, Y) \sim \mathrm{Exp}(\lambda + \mu) \tag{30}$$

10 Normal distribution

A variable $X \in \mathbb{R}$ is normal distributed for $\mu \in \mathbb{R}$ and $\sigma^2 > 0$ if

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} \cdot e^{-\frac{1}{2}x^2} \tag{31}$$

We denote:

$$X \sim \mathcal{N}(\mu, \sigma^2) \tag{32}$$

$$\mathbb{E}(X) = \mu \tag{33}$$

$$\mathrm{Var}(X) = \sigma^2 \tag{34}$$

$$\tag{35}$$

10.1 Independent random normal distributed variables

$X \sim \mathcal{N}(\mu, \sigma^2)$ and $Y \sim \mathcal{N}(\upsilon, \tau^2)$ can be merged:

$$X + Y \sim \mathcal{N}(\mu + \upsilon, \sigma^2 + \tau^2) \tag{36}$$