Elements of Statistics

Chapter 6: Random variables

Ralf Münnich, Jan Pablo Burgard and Florian Ertz

University of Trier
Faculty IV
Economic and Social Statistics Department

Winter term 2021/22

Notes

Example 6.1: Triple coin toss (1)

Triple coin toss (H: Heads, T: Tails):

Let X = Number of heads be a random variable with $x \in \{0, 1, 2, 3\}$.

Toss	Result			Realisation			Number of heads
1	H	H	H	1	1	1	3
2	H	H	T	1	1	0	2
3	H	T	H	1	0	1	2
4	H	T	T	1	0	0	1
5	T	H	H	0	1	1	2
6	T	H	T	0	1	0	1
7	T	T	H	0	0	1	1
8	T	T	T	0	0	0	0

Notes

Example 6.1: Triple coin toss (2)

Construction of the table in R:

```
result <- expand.grid(lapply(
                        X = 1:3,
                        FUN = function(x) c("H", "T")))
domain <- expand.grid(lapply(
                           X = 1:3,
                           FUN = function(x) c(1, 0)))
number_of_hats <- rowSums(domain)

Example6_1<- data.frame(result,domain,number_of_hats)
names(Example6_1)<- c("result","","","domain","","",
                        "number_of_hats")

save(Example6_1, file="Example6-1.RData")
head(Example6_1, n = 4)
```

```
      result       domain     number_of_hats
1      H H H        1 1 1            3
2      T H H        0 1 1            2
3      H T H        1 0 1            2
4      T T H        0 0 1            1
```


Notes

Definition of a random variable

Definition 6.1:
Let a probability space $(\Omega; \mathcal{S}; P)$ be given. A function

$$X : \Omega \to \mathbb{R}; \qquad \omega \mapsto X(\omega)$$

is called *random variable*, if the set

$$X^{-1}((-\infty, x]) = \{\omega \in \Omega | X(\omega) \le x\}$$

belongs to the sigma algebra $\mathcal{S}$ over Ω for all $x \in \mathbb{R}$.

Notes

Example 6.2: see Ex. 6.1 (1)

Wirtschafts- und Sozialstatistik Universität Trier

$$P(X=1) = P(\{(H,T,T),(T,H,T),(T,T,H)\}) \\ = \frac{3}{8}$$

```
load("Example6-1.RData")
sum(Example6_1$number_of_hats == 1) /
length(Example6_1$number_of_hats)
```

```
[1] 0.375
```

$$P(X\leq 1) = P(X=0)+P(X=1) \\ = \frac{1}{8}+\frac{3}{8} = \frac{1}{2}$$

```
sum(Example6_1$number_of_hats==0 |
    Example6_1$number_of_hats==1) /
    length(Example6_1$number_of_hats)
```

```
[1] 0.5
```


Notes

Example 6.2: see Ex. 6.1 (2)

Wirtschafts- und Sozialstatistik Universität Trier

$$P(X>1) = 1 - P(X\leq 1) = \frac{1}{2}$$

```
sum(Example6_1$number_of_hats > 1) /
length(Example6_1$number_of_hats)
```

```
[1] 0.5
```

$$P(0<X\leq 2) = P(X=1)+P(X=2) = \frac{3}{8}+\frac{3}{8} = \frac{3}{4}$$

```
sum(Example6_1$number_of_hats>0&
Example6_1$number_of_hats<=2) /
length(Example6_1$number_of_hats)
```

```
[1] 0.75
```


Notes

Distribution function

Wirtschafts- und Sozialstatistik Universität Trier

Definition 6.2:
The function $F(x) := P(\{X\leq x\})$, which assigns to each $x\in\mathbb{R}$ the probability that the random variable X is less than or equal to x, is called *distribution function* of X.

We use the short hand $P(X\leq x)$.

Notes

Example 6.3: see Ex. 6.2 (1)

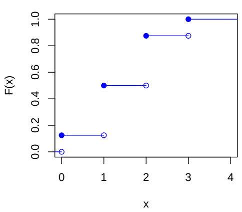
Wirtschafts- und Sozialstatistik Universität Trier

For the random variable X we have:

1. For $x<0$: $P(X\leq x)=0$
2. For $0\leq x<1$: $P(X\leq x)=P(X=0)=1/8$
3. For $1\leq x<2$: $P(X\leq x)=P(X=0)+P(X=1)=1/2$
4. For $2\leq x<3$: $P(X\leq x)=1-P(X=3)=7/8$
5. For $x\geq 3$: $P(X\leq x)=1$

Therefore, we have:

$$F(x)=\begin{cases}0 & \text{for } x<0\\ 1/8 & \text{for } 0\leq x<1\\ 1/2 & \text{for } 1\leq x<2\\ 7/8 & \text{for } 2\leq x<3\\ 1 & \text{for } x\geq 3\end{cases}$$



Notes

Example 6.3: see Ex. 6.2 (2)

Wirtschafts- und Sozialstatistik
Universität Trier

Calculation of $F(x)$ and construction of the graphic in R:

```
load("Example6-1.RData")

x6_3 <- Example6_1$number_of_hats
F_x6_3 <- cumsum(prop.table(x = table(x = x6_3)))

plot(ecdf(x6_3), col = "blue", xlab = "x", ylab = "F(x)",
     main = "", xlim = c(0, 4))
points(0:3, c(0, F_x6_3[-4]), col = "blue")

F_x6_3
```

```
    0     1     2     3
0.125 0.500 0.875 1.000
```


Notes

Properties of distribution functions

Wirtschafts- und Sozialstatistik
Universität Trier

We have:

1. $0 \leq F(x) \leq 1$ for all $x \in \mathbb{R}$
2. $\lim\limits_{x \to \infty} F(x) = 1$
3. $\lim\limits_{x \to -\infty} F(x) = 0$
4. F is monotonously increasing.
5. F has no more than a countable number of jump discontinuities.
6. F is right-continuous.

Notes

Discrete random variables

Wirtschafts- und Sozialstatistik
Universität Trier

Definition 6.3:
A random variable X is called *discrete* if it cannot take more than a countable number of values (realisations) with a positive probability.
If $x_1, \ldots, x_i$ are the realisations of X, then the probabilities $P(X = x_1), \ldots, P(X = x_i)$ contain the complete information about this random variable.

Definition 6.4:
The function $f(x)$, which is defined for all real x and given by

$$f(x) = \begin{cases} P(X = x) & \text{for all possible realisations of } X \\ 0 & \text{else} \end{cases},$$

is called *probability function* of the (discrete) random variable X.

Notes

Example 6.4: Red and blue balls (1)

Wirtschafts- und Sozialstatistik
Universität Trier

An urn contains 3 red and 7 black balls. 3 balls are drawn with replacement. Determine the probability table for the number of red balls drawn.

Furthermore, make suitable plots for the respective probability function and distribution function.

Notes

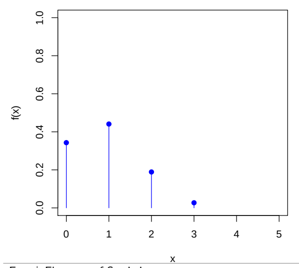
Example 6.4: Red and blue balls (2)

Wirtschafts- und Sozialstatistik Universität Trier

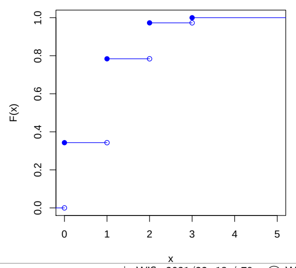
x_i	0	1	2	3
$f(x_i)$	0.343	0.441	0.189	0.027

```
x6_4 <- 0:3
f_x6_4 <- c(0.343,0.441,0.189,0.027)
F_x6_4 <- cumsum(f_x6_4)
```

Probability function



Distribution function



Notes

Example 6.4: Red and blue balls (3)

Wirtschafts- und Sozialstatistik Universität Trier

Construction of the graphics in R:

```
plot(x = x6_4, y = f_x6_4, type = "h", lwd = 3,
     xlim = c(0,5), ylim = c(0,0.6), col = "blue", xlab="x",
     ylab = "f(x)", main = "Probability function")
points(x = x6_4, y = f_x6_4, col = "blue", pch = 19)

x_axis <- c(-1,sort(x6_4),4)
y_axis <- c(0,F_x6_4,1)

plot(x = c(0,4), y = c(0,1), main = "Distribution function",
     type = "n", col = "blue", xlab = "x", ylab = "F(x)")

lines(x = x_axis[1:2],y=rep(y_axis[1],2),col="blue",lwd=2)
lines(x = x_axis[2:3],y=rep(y_axis[2],2),col="blue",lwd=2)
lines(x = x_axis[3:4],y=rep(y_axis[3],2),col="blue",lwd=2)
lines(x = x_axis[4:5],y=rep(y_axis[4],2),col="blue",lwd=2)
lines(x = x_axis[5:6],y=rep(y_axis[5],2),col="blue",lwd=2)
points(x = x_axis[1:5],y=y_axis[1:5], col="blue", pch=19)
points(x = x_axis[2:5],y=y_axis[1:4], col="blue", pch=1)
```


Notes

Example 6.5: Another random variable (1)

Wirtschafts- und Sozialstatistik Universität Trier

Let a discrete random variable have the following probability and distribution function, respectively:

$$f(x) = \begin{cases} 0.2 \cdot 0.8^x & \text{for } x = 0, 1, 2, \ldots \\ 0 & \text{else} \end{cases}$$

$$F(x) = \begin{cases} 1 - 0.8^{x+1} & \text{for } x \geq 0 \\ 0 & \text{else} \end{cases}$$

Definition of $f(x)$ and $F(x)$ in R:

```
x6_5 <- 0:10

# ATTENTION: here functions
f_x <- function(x) {0.2 * 0.8^x}

F_x <- function(x) {1 - 0.8^(x+1)}
```


Notes

Example 6.5: Another random variable (2)

Wirtschafts- und Sozialstatistik Universität Trier

We get the following tabulated results:

x_i	0	1	2	3	4	5	6	7	8	9	10
$f(x_i)$	0.200	0.160	0.128	0.102	0.082	0.066	0.052	0.042	0.034	0.027	0.021
$F(x_i)$	0.200	0.360	0.488	0.590	0.672	0.738	0.790	0.832	0.866	0.893	0.914

```
round(f_x(x6_5), digits = 3)
```

```
[1] 0.200 0.160 0.128 0.102 0.082 0.066 0.052 0.042 0.034 0.027 0.021
```

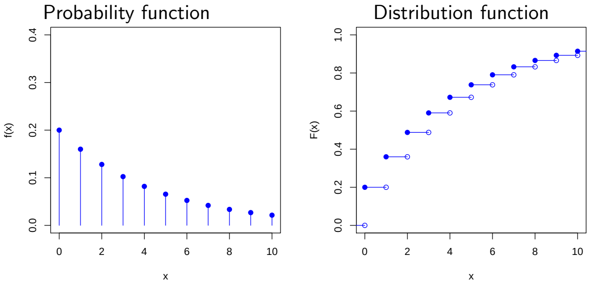
```
round(F_x(x6_5), digits = 3)
```

```
[1] 0.200 0.360 0.488 0.590 0.672 0.738 0.790 0.832 0.866 0.893 0.914
```


Notes

Example 6.5: Another random variable (3)

Wirtschafts- und Sozialstatistik
Universität Trier



The random variable has a countably infinite number of realisations.
Here, we are dealing with a geometric distribution with parameter $p = 0.2$.

Notes

Continuous random variables

Wirtschafts- und Sozialstatistik
Universität Trier

Definition 6.5:
If there is a non-negative function $f(x)$ for a random variable X, in such a way that the distribution function for all x can be described by

$$F(x) = \int_{-\infty}^{x} f(y)\, dy,$$

we call X a *continuous random variable*.

Definition 6.6:
The function $f(x)$ of Definition 7.5 is called the *density function* of the continuous random variable X.

Notes

Properties of continuous random variables

Wirtschafts- und Sozialstatistik
Universität Trier

1. The area between the density curve and the abscissa has to sum up to 1. Notice the analogy to the empirical relative frequency distribution (histogram).
2. The probability $F(x)$ that X takes on a value which is less than or equal to x is expressed in terms of the measure of the area between the density curve and the abscissa on the interval $(-\infty, x]$. Notice the analogy to the empirical distribution function.
3. $P(x_1 < X \leq x_2) = F(x_2) - F(x_1)$

Notes

Properties of continuous random variables (ctd.)

Wirtschafts- und Sozialstatistik
Universität Trier

4. $P(X = x) = 0$
   Density values cannot be interpreted as probabilities!
5. X continuous $\Rightarrow$
   $P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) =$
   $P(x_1 \leq X \leq x_2) = P(x_1 < X < x_2)$
6. Interpretation of densities: $P(x_1 < X \leq x_2) \approx f(x) \cdot \underbrace{(x_2 - x_1)}_{\text{small}}$
7. $f(x) > 1$ is possible!
8. $F'(x) = f(x)$ for all x, for which F is differentiable.

Notes

Example 6.6: A continuous random variable

Wirtschafts- und Sozialstatistik
Universität Trier

Let the continuous random variable X have the following density function:

$$f(x) = \begin{cases} 0.5 & \text{for } 1 \leq x \leq 3 \\ 0 & \text{else} \end{cases} .$$

Then, we get the distribution function

$$F(x) = \begin{cases} 0 & \text{for } x < 1 \\ 0.5x - 0.5 & \text{for } 1 \leq x \leq 3 \\ 1 & \text{for } x > 3 \end{cases} .$$

Application of $f(x)$ and $F(x)$ in R:

```
# Distinction from the latest functions
f_x6_6 <- function(x) {0.5}

F_x6_6 <- function(x) {0.5 * x - 0.5}
```


Notes

Example 6.7: An exponentially distributed RV

Wirtschafts- und Sozialstatistik
Universität Trier

Let the continuous random variable X have the following distribution function:

$$F(x) = \begin{cases} 1 - \mathrm{e}^{-\frac{1}{2}x} & \text{for } x \geq 0 \\ 0 & \text{else} \end{cases}$$

(exponential distribution with parameter $\lambda = \frac{1}{2}$). Then, differentiation yields the density function

$$f(x) = \begin{cases} \frac{1}{2}\mathrm{e}^{-\frac{1}{2}x} & \text{for } x \geq 0 \\ 0 & \text{else} \end{cases} .$$

Application of $F(x)$ and $f(x)$ in R:

```
F_x6_7 <- function(x) {1 - exp(-1/2 * x)}

f_x6_7 <- function(x) {1/2 * exp(-1/2 * x)}
```


Notes

Definition 6.7:

Wirtschafts- und Sozialstatistik
Universität Trier

Let the random variable X have the following probability or density function $f(x)$, respectively. If

$$\sum_i |f(x_i) \cdot x_i| < \infty \qquad \text{or} \qquad \int_{-\infty}^{\infty} |f(x) \cdot x| \, dx < \infty$$

holds, then

$$\mathrm{E}(X) := \sum_i f(x_i) \cdot x_i \text{ or } \mathrm{E}(X) := \int_{-\infty}^{\infty} f(x) \cdot x \, dx$$

is called the *expected value* of the discrete or continuous random variable X, respectively.

Notes

Definition 6.8:

Wirtschafts- und Sozialstatistik
Universität Trier

Let the random variable X have the following probability or density function $f(x)$, respectively. If

$$\sum_i |f(x_i) \cdot x_i^2| < \infty \qquad \text{or} \qquad \int_{-\infty}^{\infty} |f(x) \cdot x^2| \, dx < \infty$$

holds, then

$$\mathrm{Var}(X) := \sum_i (x_i - \mathrm{E}\,X)^2 \cdot f(x_i) \text{ or}$$

$$\mathrm{Var}(X) := \int_{-\infty}^{\infty} (x - \mathrm{E}\,X)^2 \cdot f(x) \, dx$$

is called the *variance* of the discrete or continuous random variable X, respectively.

Notes

Wirtschafts- und Sozialstatistik Universität Trier

Example 6.8: see Ex. 6.4

$$\begin{aligned} \mathrm{E}(X) &= \sum_{x=0}^{3} x \cdot f(x) \\ &= 0 \cdot 0.343 + 1 \cdot 0.441 + 2 \cdot 0.189 + 3 \cdot 0.027 \\ &= 0.9 \end{aligned}$$

$$\begin{aligned} \mathrm{Var}(X) &= \sum_{x=0}^{3} x^2 \cdot f(x) - \mathrm{E}(X)^2 \\ &= 0^2 \cdot 0.343 + 1^2 \cdot 0.441 + 2^2 \cdot 0.189 + 3^2 \cdot 0.027 - 0.9^2 \\ &= 1.44 - 0.9^2 = 0.63 \end{aligned}$$

Calculation of $\mathrm{E}(X)$ and $\mathrm{Var}(X)$ in R:

```
Mean_X6_8 <- weighted.mean(x = x6_4, w = f_x6_4)

Var_X6_8 <- sum(f_x6_4*(x6_4 - Mean_X6_8)^2)
```

```
Mean_X6_8
```

```
[1] 0.9
```

```
Var_X6_8
```

```
[1] 0.63
```


Notes

Wirtschafts- und Sozialstatistik Universität Trier

Example 6.9: see Ex. 6.6 (1)

$$\begin{aligned} \mathrm{E}(X) &= \int_{-\infty}^{\infty} x \cdot f(x)dx = \int_{1}^{3} x \cdot 0.5dx \\ &= \left[\frac{1}{2} \cdot x^2 \cdot 0.5\right]_1^3 = \left[\frac{1}{4} \cdot x^2\right]_1^3 \\ &= \frac{9}{4} - \frac{1}{4} = 2 \end{aligned}$$

Calculation of $\mathrm{E}(X)$ in R:

```
Mean_X6_9 <- integrate(f = function(x){0.5*x}, lower = 1,
                       upper = 3)$value
Mean_X6_9
```

```
[1] 2
```


Notes

Wirtschafts- und Sozialstatistik Universität Trier

Example 6.9: see Ex. 6.6 (2)

$$\begin{aligned} \mathrm{Var}(X) &= \int_{-\infty}^{\infty} x^2 \cdot f(x)dx - \mathrm{E}(X)^2 = \int_{1}^{3} x^2 \cdot 0.5dx - 2^2 \\ &= \left[\frac{1}{6}x^3\right]_1^3 - 4 \\ &= \frac{27}{6} - \frac{1}{6} - 4 = \frac{1}{3} \end{aligned}$$

Calculation of $\mathrm{Var}(X)$ in R:

```
Var_X6_9 <- integrate(f = function(x){0.5*x^2}, lower = 1,
                      upper = 3
                      )$value - Mean_X6_9^2
Var_X6_9
```

```
[1] 0.3333333
```


Notes

Wirtschafts- und Sozialstatistik Universität Trier

Linear transformation of a random variable

Let $Y = a + b \cdot X$, then

$$\mathrm{E}(Y) = a + b \cdot \mathrm{E}(X)$$

$$\mathrm{Var}(Y) = b^2 \cdot \mathrm{Var}(X)$$

Example 6.10:

X: Filling weight of a package of detergent in kg
Y: Deviation from targeted weight of 5 kg in g
Then $Y = (X - 5) \cdot 1000 = -5000 + 1000 \cdot X$ and therefore $a = -5000$ and $b = 1000$.

Example 6.11: (see Example 6.4)

Now we are interested in the share of black balls
(earlier: number of red balls):
$Y = \dfrac{n - X}{n} = 1 - \frac{1}{n} \cdot X$ with $n = 3$

Notes

Standard transformation

Wirtschafts- und Sozialstatistik Universität Trier

The special linear transformation

$$Z = \frac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}} = -\frac{\mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}} + \frac{1}{\sqrt{\mathrm{Var}(X)}} \cdot X$$

is called standard transformation of random variable X (see Chapter 4).

We have $\mathrm{E}(Z) = 0$ and $\mathrm{Var}(Z) = 1$.

Notes

Example 6.12: see Ex. 6.4 and 6.8

Wirtschafts- und Sozialstatistik Universität Trier

X	0	1	2	3
Z	$-0.9/\sqrt{0.63}$	$0.1/\sqrt{0.63}$	$1.1/\sqrt{0.63}$	$2.1/\sqrt{0.63}$
$f(x)$	0.343	0.441	0.189	0.027

Calculation of Z in R:

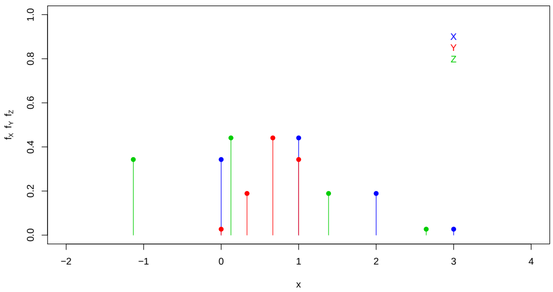
```
Z6_12 <- (x6_4 - Mean_X6_8) / sqrt(Var_X6_8)
round(Z6_12, digits = 3)
```

```
[1] -1.134  0.126  1.386  2.646
```


Notes

Visualisation for Examples 6.11 and 6.12

Wirtschafts- und Sozialstatistik Universität Trier

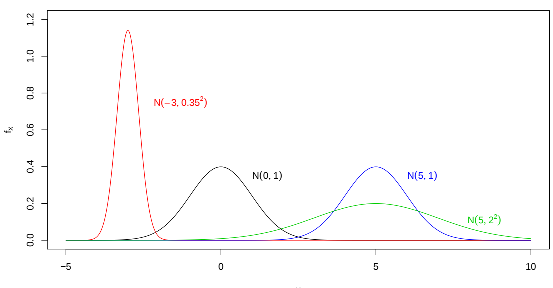


$$Y = \frac{3 - X}{3} \qquad Z = \frac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var}(X)}}$$

Notes

Different normal distributions $N(\mu; \sigma^2)$

Wirtschafts- und Sozialstatistik Universität Trier



$$f(x) = \varphi(x \mid \mu; \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

Notes

Quantiles of distributions

Wirtschafts- und Sozialstatistik Universität Trier

Definition of a quantile (see Schaich and Münnich, 2001):

For a random variable X, a value x, which satisfies the inequalities

$$P(X \leq x) \geq p \qquad \text{and} \qquad P(X \geq x) \leq 1 - p$$

for $0 < p < 1$, is called its *quantile of order p* (p-quantile).

The median $x_{0.5}$ (also called the 0.5-quantile), as well as the first and third quartile ($p = 0.25$ and $p = 0.75$, respectively) are particularly interesting.

For continuous random variables (with a strictly monotonous distribution function) the p-quantile equals $x_p = F^{-1}(p)$.

Schaich, E. and Münnich, R. (2001): Mathematische Statistik für Ökonomen: Lehrbuch. Vahlen.

Notes

Example 6.13: Quantiles of exp. distr. (1)

Wirtschafts- und Sozialstatistik Universität Trier

Let the random variable X follow an exponential distribution with parameter $\lambda = \frac{1}{2}$. The distribution function is:

$$F(x) = \begin{cases} 1 - e^{-\frac{1}{2}\cdot x} & \text{for } x \geq 0 \\ 0 & \text{else} \end{cases} \quad .$$

The p-quantile is derived as follows:

$$\begin{aligned} p &= 1 - e^{-\frac{1}{2}\cdot x} && \big| -p \quad \big| + e^{-\frac{1}{2}\cdot x} \\ e^{-\frac{1}{2}\cdot x} &= 1 - p && \big| \ln \\ -0.5 \cdot x &= \ln(1-p) && \big| : (-0.5) \\ x &= -2\ln(1-p) \end{aligned}$$

Notes

Example 6.13: Quantiles of exp. distr. (2)

Wirtschafts- und Sozialstatistik Universität Trier

Therefore, the median is

$$\begin{aligned} x_{0.5} &= -2\ln(1-0.5) = -2\ln\left(\frac{1}{2}\right) = -2(\ln 1 - \ln 2) \\ &= -2\ln 1 + 2\ln 2 = 2\ln 2 \approx 1.3863 \end{aligned}$$

and the first quantile is
$x_{0.25} = -2\ln\frac{3}{4} \approx 0.5754$.

Calculation of $x_{0.5}$ and $x_{0.25}$ in R:

```
q_050 <- -2 * log(1 - 0.5)
q_025 <- -2 * log(1 - 0.25)
```

```
round(q_050, digits = 4)
[1] 1.3863
```

```
round(q_025, digits = 4)
[1] 0.5754
```


Notes

Markov's and Tchebysheff's inequality

Wirtschafts- und Sozialstatistik Universität Trier

Theorem 6.1 (Markov's inequality): If a random variable X only takes on non-negative values and the expected value $\mathrm{E}(X)$ exists, the following approximation holds for every $x^* > 0$:

$$P(X \geq x^*) \leq \frac{\mathrm{E}(X)}{x^*} \quad .$$

Theorem 6.2 (Tchebysheff's inequality): If the variance $\mathrm{Var}(X)$ of a random variable X exists, the following holds for $\varepsilon > 0$:

$$P(|X - \mathrm{E}(X)| \geq \varepsilon) \leq \frac{\mathrm{Var}(X)}{\varepsilon^2} \quad .$$

Notice the special case where $\varepsilon = k \cdot \sqrt{\mathrm{Var}(X)}$.

Notes

Example 6.14: An inequality

Wirtschafts- und Sozialstatistik
Universität Trier

Let a non-negative random variable X have the expected value $\mathrm{E}(X) = 10$ (applies for discrete as well as continuous variables).

We can approximate:

$$P(X \geq 25) \leq \frac{10}{25} = 0.4$$

$$P(X \geq 40) \leq \frac{10}{40} = 0.25$$

$$P(X \geq 5) \leq \frac{10}{5} = 2\,.$$

The third row is a *trivial* approximation as probabilities are bounded by 0 and 1.

Notes

Example 6.15: Another inequality

Wirtschafts- und Sozialstatistik
Universität Trier

Let the expected value $\mathrm{E}(X) = 2$ and the variance $\mathrm{Var}(X) = 36$ of a random variable X be known.

Then we have:

$$P(-8 < X < 12) = P(|X - 2| < 10) \geq 1 - \frac{36}{100} = 0.64$$

$$P(|X - 2| \geq 10) \leq \frac{36}{100} = 0.36$$

$$P(X \leq -3 \vee X \geq 7) = P(|X - 2| \geq 5) \leq \frac{36}{25} = 1.44\,.$$

Notes

Example 6.16: More dimensions (1)

Wirtschafts- und Sozialstatistik
Universität Trier

We are looking at two- and multi-dimensional random variables.

a)
Random questioning of a person with replacement (income; age):
The resulting observation is $(1815; 25)$.

b)
Two rolls of a dice:
The resulting pair of number of pips is $(4; 6)$.
We could as well be interested in the overall sum of pips or the product of the number of pips $(10; 24)$.

Notes

Example 6.16: More dimensions (2)

Wirtschafts- und Sozialstatistik
Universität Trier

c) An urn contains $N = 100$ balls, of which 30 are red (r), 20 are white (w) and 50 are black (s). How does the sample space of this experiment look like, if we draw 3 balls? $X =$ (number r, number w)

ω	$X(\omega)$	ω	$X(\omega)$	ω	$X(\omega)$
(s,s,s)	(0,0)	(r,s,s)	(1,0)	(w,s,s)	(0,1)
(s,s,r)	(1,0)	(r,s,r)	(2,0)	(w,s,r)	(1,1)
(s,s,w)	(0,1)	(r,s,w)	(1,1)	(w,s,w)	(0,2)
(s,r,s)	(1,0)	(r,r,s)	(2,0)	(w,r,s)	(1,1)
(s,r,r)	(2,0)	(r,r,r)	(3,0)	(w,r,r)	(2,1)
(s,r,w)	(1,1)	(r,r,w)	(2,1)	(w,r,w)	(1,2)
(s,w,s)	(0,1)	(r,w,s)	(1,1)	(w,w,s)	(0,2)
(s,w,r)	(1,1)	(r,w,r)	(2,1)	(w,w,r)	(1,2)
(s,w,w)	(0,2)	(r,w,w)	(1,2)	(w,w,w)	(0,3)

Notes

Example 6.16: More dimensions (3)

Wirtschafts- und Sozialstatistik Universität Trier

Construction of the table in R:

```
omega <- expand.grid(lapply(X = 1:3,
                     FUN = function(x) c("s", "r", "w")))
Number_of_r <- rowSums(omega == "r")
Number_of_w <- rowSums(omega == "w")
X6_16 <- cbind(Number_of_r, Number_of_w)

Example6_16 <- data.frame(omega, X6_16)
names(Example6_16) <- c("Omega","","",
                        "Number_of_r","Number_of_w")

head(Example6_16)
```

```
  Omega Number_of_r Number_of_w
1 s s s       0           0
2 r s s       1           0
3 w s s       0           1
4 s r s       1           0
5 r r s       2           0
6 w r s       1           1
```


Notes

Example 6.16: More dimensions (4)

Wirtschafts- und Sozialstatistik Universität Trier

The set of all possible realisations is

$$\{(x_1, x_2) \,|\, x_1, x_2 \in \mathbb{N}_0, 0 \leq x_1 + x_2 \leq 3\}\,.$$

For instance, we could determine $P(X_1 = x_1; X_2 = x_2)$, $P(X_1 \leq x_1; X_2 \leq x_2)$, $P(X_1 \leq x_1 \vee X_2 \leq x_2)$ or $P(X_1 \leq x_1 | X_2 = x_2)$.

Notes

Multi-dimensional random variables

Wirtschafts- und Sozialstatistik Universität Trier

Multi-dimensional random variables can be considered as a generalisation of one-dimensional random variables.

The inverse images of half-open n-intervals must again be part of the sigma algebra over Ω.

Notes

Distribution function of multi-dimensional random variables

Wirtschafts- und Sozialstatistik Universität Trier

The function

$$F_X(x_1, \ldots, x_n) = P(X_1 \leq x_1, X_2 \leq x_2, \ldots, X_n \leq x_n) \qquad ,$$

which gives the probability that X_1 is at most x_1 and X_n is at most x_n for all real n-tuple is called distribution function of the random vector $X_1, X_2, \ldots, X_n$.

Interval probabilities in the two-dimensional case:

$$P(x_1' < X_1 \leq x_1'', x_2' < X_2 \leq x_2'') = F(x_1'', x_2'') - F(x_1', x_2'') - F(x_1'', x_2') + F(x_1', x_2') \quad .$$

Notes

Discrete random vectors

Wirtschafts- und Sozialstatistik Universität Trier

Discrete random vector

A random vector $\mathbf{X}$ is called a multi-dimensional discrete random variable if each of its components can take on at most a countable number of values.

Probability function of a discrete random vector

The function $f(x_1, x_2)$, which is defined for all real pairs of numbers (x_1, x_2) and which is characterised by

$$f(x_1, x_2) = \begin{cases} P(X_1 = x_{1j}, X_2 = x_{2k}) & \text{for all } j, k \\ 0 & \text{else} \end{cases}$$

is called the probability function of the discrete random vector $\mathbf{X}$.

Notes

Two-dimensional discrete random variables

Wirtschafts- und Sozialstatistik Universität Trier

$X_1 \backslash X_2$	x_{21}	$\cdots$	x_{2k}	$\cdots$	x_{2r}	$\sum$
x_{11}	$f(x_{11}, x_{21})$	$\cdots$	$f(x_{11}, x_{2k})$	$\cdots$	$f(x_{11}, x_{2r})$	$f_{X_1}(x_{11})$
$\vdots$	$\vdots$	$\ddots$	$\vdots$		$\vdots$	$\vdots$
x_{1j}	$f(x_{1j}, x_{21})$	$\cdots$	$f(x_{1j}, x_{2k})$	$\cdots$	$f(x_{1j}, x_{2r})$	$f_{X_1}(x_{1j})$
$\vdots$	$\vdots$		$\vdots$	$\ddots$	$\vdots$	$\vdots$
x_{1m}	$f(x_{1m}, x_{21})$	$\cdots$	$f(x_{1m}, x_{2k})$	$\cdots$	$f(x_{1m}, x_{2r})$	$f_{X_1}(x_{1m})$
$\sum$	$f_{X_2}(x_{21})$	$\cdots$	$f_{X_2}(x_{2k})$	$\cdots$	$f_{X_2}(x_{2r})$	1

Notes

Properties of discrete random vectors

Wirtschafts- und Sozialstatistik Universität Trier

1. We have:

$$\sum_j \sum_k f(x_{1j}, x_{2k}) = 1 .$$

2. Distribution function:

$$F(x_1, x_2) = \sum_{x_{1j} \leq x_1} \sum_{x_{2k} \leq x_2} f(x_{1j}, x_{2k})$$

3. Interval probabilities:

$$P(x_1' < X_1 \leq x_1'', x_2' < X_2 \leq x_2'') = \sum_{x_1' < x_{1j} \leq x_1''} \sum_{x_2' < x_{2k} \leq x_2''} f(x_{1j}, x_{2k})$$

Notes

Marginal distributions of bivariate distributions

Wirtschafts- und Sozialstatistik Universität Trier

In addition to the joint distribution of the random vector (X_1, X_2) with the distribution function $F(x_1, x_2)$, the *marginal distributions*, ergo the univariate distributions of the random variables involved in the distribution functions $F_{X_1}(x_1)$ and $F_{X_2}(x_2)$, may be considered as well. We obtain those by

$$F_{X_1}(x_1) = \sum_{x_{1j} \leq x_1} \sum_k f(x_{1j}, x_{2k}) \quad \text{or} \quad F_{X_2}(x_2) = \sum_j \sum_{x_{2k} \leq x_2} f(x_{1j}, x_{2k}) \quad ,$$

and thus by adding up all probabilities of the variable which is not of interest.

The indexation of the marginal distribution functions is used for unique identification.

Notes

Example 6.17: see Ex. 6.16 c) (1)

We are interested in the random vector (number of red balls, number of white balls). For example, we have:

$$f(1,0) = 3 \cdot 0.3^1 \cdot 0.2^0 \cdot 0.5^2 = 0.225 \quad .$$

Calculation of $f(1,0)$ in R:

```
Number_of_s <- rowSums(omega == "s")
Example6_16 <- cbind(Example6_16, Number_of_s)
Probs <- 0.3^Example6_16$Number_of_r *
           0.2^Example6_16$Number_of_w *
           0.5^Example6_16$Number_of_s

Example6_16 <- cbind(Example6_16, Probs)

pos <- which(Example6_16$Number_of_r == 1 &
             Example6_16$Number_of_w == 0)
f_1_0 <- sum(Example6_16[pos, 7])
f_1_0
```

```
[1] 0.225
```


Notes

Example 6.17: see Ex. 6.16 c) (2)

Finally, we get the following probability table:

$X_1 \mid X_2$	$x_{21}=0$	$x_{22}=1$	$x_{23}=2$	$x_{24}=3$	$\sum$
$x_{11}=0$	0.125	0.150	0.060	0.008	0.343
$x_{12}=1$	0.225	0.180	0.036	0.000	0.441
$x_{13}=2$	0.135	0.054	0.000	0.000	0.189
$x_{14}=3$	0.027	0.000	0.000	0.000	0.027
$\sum$	0.512	0.384	0.096	0.008	1.000

Probability table in R:

```
X1_6_17 <- 0:3 ; X2_6_17 <- 0:3
ProbTable6_17 <- matrix(c(0.125, 0.150, 0.060, 0.008,
                          0.225, 0.180, 0.036, 0.000,
                          0.135, 0.054, 0.000, 0.000,
                          0.027, 0.000, 0.000, 0.000),
                        ncol = length(X2_6_17),
                        byrow = TRUE)
dimnames(ProbTable6_17) <- list(X1_6_17, X2_6_17)
ProbTable_new6_17 <- addmargins(ProbTable6_17)
```


Notes

Example 6.17: see Ex. 6.16 c) (3)


```
ProbTable_new6_17
```

```
        0     1     2     3   Sum
0   0.125 0.150 0.060 0.008 0.343
1   0.225 0.180 0.036 0.000 0.441
2   0.135 0.054 0.000 0.000 0.189
3   0.027 0.000 0.000 0.000 0.027
Sum 0.512 0.384 0.096 0.008 1.000
```

For $F(1,2)$ we have:

$X_1 \mid X_2$	$x_{21}=0$	$x_{22}=1$	$x_{23}=2$	$x_{24}=3$	$\sum$
$x_{11}=0$	0.125	0.150	0.060	0.008	0.343
$x_{12}=1$	0.225	0.180	0.036	0.000	0.441
$x_{13}=2$	0,135	0.054	0.000	0.000	0.189
$x_{14}=3$	0.027	0.000	0.000	0.000	0.027
$\sum$	0.512	0.384	0.096	0.008	1.000

$$F(1,2) = 0.125 + 0.150 + 0.060 + 0.225 + 0.180 + 0.036 = 0.776$$

Notes

Example 6.17: see Ex. 6.16 c) (4)

Determination of the joint distribution function in R:

```
F_x1_x2 <- t(apply(X = apply(X = ProbTable6_17,
                             MARGIN = 2, FUN = cumsum),
                   MARGIN = 1, FUN = cumsum))
F_x1_x2
```

```
      0     1     2     3
0 0.125 0.275 0.335 0.343
1 0.350 0.680 0.776 0.784
2 0.485 0.869 0.965 0.973
3 0.512 0.896 0.992 1.000
```

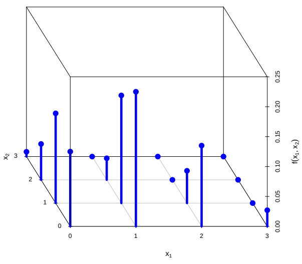
```
F_x1_x2[rownames(F_x1_x2) == 1, colnames(F_x1_x2) == 2]
```

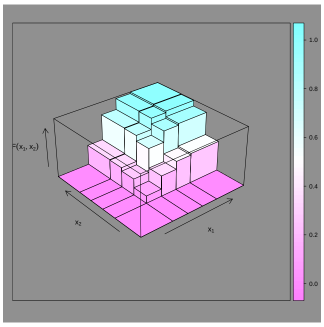
```
[1] 0.776
```


Notes

Example 6.17: see Ex. 6.16 c) (5)

Wirtschafts- und Sozialstatistik
Universität Trier





Notes

Continuous random vectors

Wirtschafts- und Sozialstatistik
Universität Trier

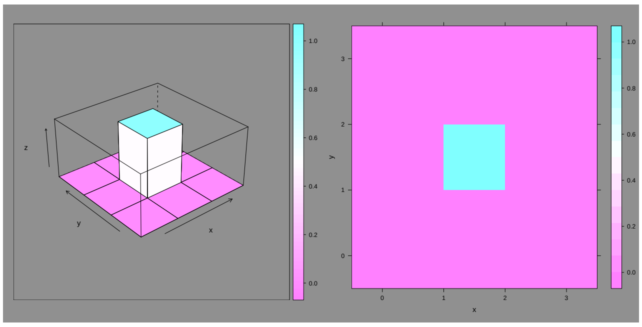
Continuous random vectors are defined analogously to continuous random variables. We have:

1. $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x_1, x_2)\, dx_2 dx_1 = 1$
2. $F(x_1, x_2) = \int_{-\infty}^{x_1}\int_{-\infty}^{x_2} f(y_1, y_2)\, dy_2 dy_1$
3. $P(x_1' < X_1 \leq x_1'', x_2' < X_2 \leq x_2'') = \int_{x_1'}^{x_1''}\int_{x_2'}^{x_2''} f(x_1, x_2)\, dx_2 dx_1$
4. $f(x_1, x_2) = \dfrac{\partial^2 F(x_1, x_2)}{\partial x_1\, \partial x_2}$ (assuming differentiability)
5. Marg. distr.: $f(x_1) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_2$ and $f(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1$

Notes

Rectangular distribution

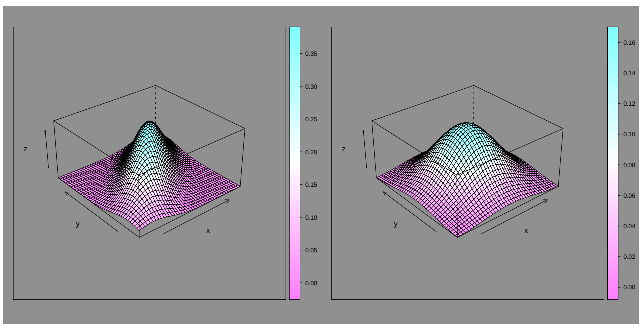
Wirtschafts- und Sozialstatistik
Universität Trier



Notes

Bivariate normal distribution

Wirtschafts- und Sozialstatistik
Universität Trier



Notes

Example 6.18: A continuous random vector (1)

Wirtschafts- und Sozialstatistik
Universität Trier

Let the density function of a continuous random vector be

$$f(x_1, x_2) = \begin{cases} 6 \cdot \exp(-2x_1) \cdot \exp(-3x_2) & \text{for } x_1, x_2 > 0 \\ 0 & \text{else} \end{cases} .$$

Using integration we get the following distribution function:

$$F(x_1, x_2) = \begin{cases} (1 - \exp(-2x_1)) \cdot (1 - \exp(-3x_2)) & \text{for } x_1, x_2 > 0 \\ 0 & \text{else} \end{cases} .$$

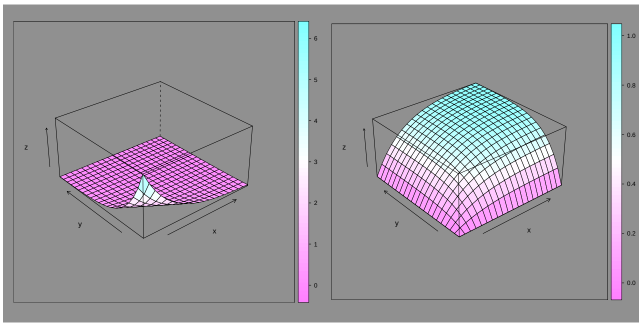
For $x_2 > 0$ we get the following marginal density function for random variable X_2:

$$f(x_2) = \int_{-\infty}^{\infty} f(x_1, x_2) dx_1 = \int_{0}^{\infty} 6 \cdot \exp(-2x_1 - 3x_2) dx_1 = 3 \cdot \exp(-3x_2) \quad .$$

Notes

Example 6.18: A continuous random vector (2)

Wirtschafts- und Sozialstatistik
Universität Trier



Notes

Stochastical independence

Wirtschafts- und Sozialstatistik
Universität Trier

Let $F(x_1, x_2)$ be the joint distribution function of the random vector $\mathbf{X}$ and let $F(x_1)$ and $F(x_2)$ be the marginal distribution functions. Two random variables X_1 and X_2 are called stochastically independent if and only if we have

$$F(x_1, x_2) = F(x_1) \cdot F(x_2)$$

for all $(x_1, x_2) \in \mathbb{R}$. Otherwise, they are called stochastically dependent. Stochastical independence may be proven using probabilities or probability functions and density functions as well (see Schaich and Münnich, 2001).

Notes

Example 6.19: see Ex. 6.17

Wirtschafts- und Sozialstatistik
Universität Trier

$f_{X_1}(0) \cdot f_{X_2}(0) = 0.343 \cdot 0.512 = 0.175616 \neq 0.125 = f(0,0)$

X_1 and X_2 are stochastically dependent.

Checking in R:

```
(ProbTable_new6_17[1,5] * ProbTable_new6_17[5,1]) ==
ProbTable_new6_17[1,1]
```

```
[1] FALSE
```

Example 6.20 (1)

Let the random vector $\mathbf{X}$ have the following probability table:

$X_1 \mid X_2$	2	4	6	$\sum$
1	0.05	0.14	0.01	0.20
5	0.20	0.56	0.04	0.80
$\sum$	0.25	0.70	0.05	1.00

Notes

Example 6.20 (2)

Probability table in R:

```
ProbTable6_20 <- matrix(c(0.05,0.14,0.01,
                          0.20,0.56,0.04),
                        ncol = 3, byrow = TRUE)

X1_6_20 <- c(1, 5)
X2_6_20 <- seq(2, 6, 2)

rownames(ProbTable6_20) <- X1_6_20
colnames(ProbTable6_20) <- X2_6_20

ProbTable6_20 <- addmargins(ProbTable6_20)

ProbTable6_20
```

```
       2    4    6 Sum
1   0.05 0.14 0.01 0.2
5   0.20 0.56 0.04 0.8
Sum 0.25 0.70 0.05 1.0
```


Notes

Example 6.20 (3)

We have $f_{X_1}(x_{1i}) \cdot f_{X_2}(x_{2j}) = f(x_{1i}, x_{2j})$ for all i, j. Therefore, X_1 and X_2 are stochastically independent.

Checking of stochastic independence in R:

```
round(ProbTable6_20[3,] * ProbTable6_20[1,4],4) ==
ProbTable6_20[1,]
```

```
   2     4     6   Sum
TRUE  TRUE  TRUE  TRUE
```

```
round(ProbTable6_20[3,] * ProbTable6_20[2,4],4) ==
ProbTable6_20[2,]
```

```
   2     4     6   Sum
TRUE  TRUE  TRUE  TRUE
```


Notes

Covariance of two random variables

The covariance of two random variables X_1 and X_2 is defined as:

$$\operatorname{Cov}(X_1, X_2) = \operatorname{E}\left((x_1 - \operatorname{E} X_1) \cdot (x_2 - \operatorname{E} X_2)\right).$$

For discrete random variables we have:

$$\operatorname{Cov}(X_1, X_2) = \sum_i \sum_j (x_{1i} - \operatorname{E} X_1) \cdot (x_{2j} - \operatorname{E} X_2) \cdot f(x_{1i}, x_{2j}).$$

Analoguously, for continuous random variables we have:

$$\operatorname{Cov}(X_1, X_2) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} (x_1 - \operatorname{E} X_1) \cdot (x_2 - \operatorname{E} X_2) f(x_1, x_2) dx_1 dx_2 .$$

Furthermore, the displacement law holds:

$$\operatorname{Cov}(X_1, X_2) = \operatorname{E}(X_1 \cdot X_2) - \operatorname{E} X_1 \cdot \operatorname{E} X_2 .$$

Notes

Example 6.21: see Ex. 6.17 (1)

$$\operatorname{E} X_1 = 0 \cdot 0.343 + 1 \cdot 0.441 + 2 \cdot 0.189 + 3 \cdot 0.027 = 0.9$$
$$\operatorname{E} X_2 = 0 \cdot 0.512 + 1 \cdot 0.384 + 2 \cdot 0.096 + 3 \cdot 0.008 = 0.6$$

Calculation of E X_1 and E X_2 in R:

```
f_X1_6_17 <- rowSums(ProbTable6_17)
f_X2_6_17 <- colSums(ProbTable6_17)

Mean_X1 <- sum(f_X1_6_17 * X1_6_17)
Mean_X2 <- sum(f_X2_6_17 * X2_6_17)

Mean_X1_old <- Mean_X1
Mean_X2_old <- Mean_X2
```

```
Mean_X1
[1] 0.9
```

```
Mean_X2
[1] 0.6
```


Notes

Example 6.21: see Ex. 6.17 (2)

$$\begin{aligned}\text{Cov}(X_1, X_2) &= 0\cdot 0\cdot 0.125 + 0\cdot 1\cdot 0.150 + 0\cdot 2\cdot 0.060 + 0\cdot 3\cdot 0.008\\ &\quad + 1\cdot 0\cdot 0.225 + 1\cdot 1\cdot 0.180 + 1\cdot 2\cdot 0.036 + 1\cdot 3\cdot 0.000\\ &\quad + 2\cdot 0\cdot 0.135 + 2\cdot 1\cdot 0.054 + 2\cdot 2\cdot 0.000 + 2\cdot 3\cdot 0.000\\ &\quad + 3\cdot 0\cdot 0.027 + 3\cdot 1\cdot 0.000 + 3\cdot 2\cdot 0.000 + 3\cdot 3\cdot 0.000\\ &\quad - 0.9\cdot 0.6\\ &= -0.18\end{aligned}$$

Calculation of $\text{Cov}(X_1, X_2)$ in R:

```
Intermed_matrix6_21 <- matrix(
               rep(x = X2_6_17, times = length(X1_6_17)),
               ncol = length(X2_6_17), byrow = TRUE)

Cov_X1_X2 <- sum(Intermed_matrix6_21 * X2_6_17 *
                 ProbTable6_17) - Mean_X1 * Mean_X2
Cov_X1_X2_old <- Cov_X1_X2
Cov_X1_X2
```

```
[1] -0.18
```


Notes

Example 6.22: see Ex. 6.20

$$\begin{aligned}\text{Cov}(X_1, X_2) &= 1\cdot 2\cdot 0.05 + 1\cdot 4\cdot 0.14 + 1\cdot 6\cdot 0.01 + 5\cdot 2\cdot 0.2 + 5\cdot 4\cdot 0.56\\ &\quad + 5\cdot 6\cdot 0.04 - (0.2 + 5\cdot 0.8)\cdot(2\cdot 0.25 + 4\cdot 0.7 + 6\cdot 0.05)\\ &= 15.12 - 15.12 = 0\end{aligned}$$

Notice that X_1 and X_2 are stochastically independent!

Calculation of $\text{Cov}(X_1, X_2)$ in R:

```
i <- 1:3 ; ProbTable6_20 <- ProbTable6_20[-3,-4]
Mean_X1 <- weighted.mean(x = X1_6_20,
                         w = prop.table(ProbTable6_20[,2]))
Mean_X2 <- weighted.mean(x = X2_6_20,
                         w = prop.table(ProbTable6_20[2,]))

Cov_X1_X2 <- (sum(X1_6_20[1] * X2_6_20[i] *
                  ProbTable6_20[1,i]) +
              sum(X1_6_20[2] * X2_6_20[i] *
              ProbTable6_20[2,i])) - Mean_X1 * Mean_X2
Cov_X1_X2
```

```
[1] 0
```


Notes

Independence and uncorrelatedness

If $\text{Cov}(X_1, X_2) = 0$, then the random variables X_1 and X_2 are called uncorrelated.
We have:

$$\text{Independence} \quad \begin{matrix}\Rightarrow \\ \not\Leftarrow\end{matrix} \quad \text{Uncorrelatedness}$$

Example 6.23:

X_1 | X_2	-2	0	1	$\sum$
0	0.125	0.000	0.250	0.375
1	0.125	0.250	0.250	0.625
$\sum$	0.250	0.250	0.500	1.000

We have $\text{Cov}(X_1, X_2) = 0$ but $f(0,0) \neq f_{X_1}(0)\cdot f_{X_2}(0)$ as well. Therefore, X_1 and X_2 are uncorrelated but not independent.

Notes

Correlation of two random variables

The correlation coefficient of Bravais-Pearson for two random variables X_1 and X_2 is defined as:

$$\varrho_{X_1,X_2} = \frac{\text{Cov}(X_1, X_2)}{\sqrt{\text{Var}\, X_1 \cdot \text{Var}\, X_2}}$$

Notes

Example 6.24: see Ex. 6.21

Wirtschafts- und Sozialstatistik Universität Trier

By using $\operatorname{Cov}(X_1, X_2) = -0.18$ as well as $\operatorname{Var} X_1 = 0.630$ and $\operatorname{Var} X_2 = 0.480$, we finally have:

$$\varrho_{X_1,X_2} = \frac{-0.18}{\sqrt{0.630 \cdot 0.480}} = -0.3273\,.$$

Calculation of ϱ_{X_1,X_2} in R:

```
Var_X1 <- sum(f_X1_6_17 * (X1_6_17 - Mean_X1_old)^2)

Var_X2 <- sum(f_X2_6_17 * (X2_6_17 - Mean_X2_old)^2)

Cor_X1_X2 <- Cov_X1_X2_old / (sqrt(Var_X1 * Var_X2))

round(Cor_X1_X2, digits = 4)
[1] -0.3273
```


Notes

Properties of the correlation coefficient

Wirtschafts- und Sozialstatistik Universität Trier

1. Let Z_1 and Z_2 be the standardised random variables of the random variables X_1 and X_2. We then have:
$$\varrho_{X_1,X_2} = \operatorname{Cov}(Z_1, Z_2)\,.$$
2. Generally $-1 \le \varrho_{X_1,X_2} \le 1$.
3. If $X_2 = a_0 + a_1 \cdot X_1$ and $a_1 \neq 0$, it follows that $|\varrho_{X_1,X_2}| = 1$ (where the reverse holds as well).
4. If
$$\begin{aligned} U_1 &= a_0 + a_1 \cdot X_1 \qquad (a_1 \neq 0)\\ U_2 &= b_0 + b_1 \cdot X_2 \qquad (b_1 \neq 0)\end{aligned}$$
are linear transformations of the random variables X_1 and X_2, we have
$$\varrho_{U_1,U_2} = \operatorname{sgn}(a_1 \cdot b_1) \cdot \varrho_{X_1,X_2} \qquad .$$

Notes

Example 6.25: (see Example 6.24)

Wirtschafts- und Sozialstatistik Universität Trier

X_1:	Number of red balls	Y_1:	Share of red balls
X_2:	Number of white balls	Y_2:	Share of white balls

We have $Y_1 = X_1/3$ and $Y_2 = X_2/3$. Furthermore, we already know that $\varrho_{X_1,X_2} = -0.3273$.

Finally, we get

a) $\varrho_{Y_1,Y_2} = -0.3273$,

b) $\varrho_{X_1,Y_1} = 1$,

c) $\varrho_{X_1,Y_2} = -0.3273\,.$

Notes

More than two random variables (1)

Wirtschafts- und Sozialstatistik Universität Trier

1. The n random variables $X_1, \ldots, X_n$ are called collectively stochastically independent, if
$$F(x_1, \ldots, x_n) = F(x_1) \cdot \ldots \cdot F(x_n)$$
(analoguously for density and probability functions).
2. The n random variables $X_1, \ldots, X_n$ are called pairwise stochastically independent, if for two arbitrary but different random variables X_i and X_j we have:
$$F(x_i, x_j) = F(x_i) \cdot F(x_j)$$
(analoguously for density and probability functions).
3. We have:
collectively stochastically independent $\Rightarrow$
pairwise stochastically independent $\Rightarrow$ pairwise uncorrelated

Notes

More than two random variables (2)

Wirtschafts- und Sozialstatistik Universität Trier

4. Variance-covariance matrix:

$$\Sigma = \begin{pmatrix} \operatorname{Var} X_1 & \operatorname{Cov}(X_1, X_2) & \cdots & \operatorname{Cov}(X_1, X_n) \\ \operatorname{Cov}(X_2, X_1) & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \operatorname{Cov}(X_{n-1}, X_n) \\ \operatorname{Cov}(X_n, X_1) & \cdots & \operatorname{Cov}(X_n, X_{n-1}) & \operatorname{Var} X_n \end{pmatrix}$$

If Σ is a diagonal matrix, then the n random variables are pairwise uncorrelated.

Notes

Functions of random variables

Wirtschafts- und Sozialstatistik Universität Trier

Example 6.26:

a) Two rolls of a dice: We are interested in the overall number of pips $Y = X_1 + X_2$.

b) $N = 101$ balls $(0, ..., 100)$: n balls are drawn with replacement.
$Y_1 = \frac{1}{2}(X_1 + X_2)$
$Y_2 = \frac{1}{20}(X_1 + \cdots + X_{20})$

c) Construction of cylindric components (technical QC): X_1 is the component's diameter and X_2 is its length. Then
$Y = \frac{\pi}{4} \cdot X_1^2 \cdot X_2$
is its volume.

Notes

Expected value and variance of linearly transformed random variables

Wirtschafts- und Sozialstatistik Universität Trier

If $Y = a_0 + \sum_{i=1}^{n} a_i X_i$ is a general linear transformation of n random variables, then

$$\operatorname{E} Y = a_0 + \sum_{i=1}^{n} a_i \operatorname{E} X_i$$

is the expected value of the transformed random variable Y. We call E a linear operator! Furthermore

$$\begin{aligned} \operatorname{Var} Y &= \sum_{i=1}^{n}\sum_{j=1}^{n} a_i \cdot a_j \cdot \operatorname{Cov}(X_i, X_j) \\ &= \sum_{i=1}^{n} a_i^2 \cdot \operatorname{Var} X_i + 2 \cdot \sum\sum_{i<j} a_i \cdot a_j \cdot \operatorname{Cov}(X_i, X_j) \end{aligned}$$

is the variance of the transformed random variable Y.

Notes

Example 6.27: see Ex. 6.26 b) (1)

Wirtschafts- und Sozialstatistik Universität Trier

An urn contains $N = 100$ balls $(1,...,100)$. $n = 3$ balls are drawn with replacement, where X_i is the number drawn in the i-th draw. Then we have:

$$\begin{aligned} \operatorname{E} X_i &= \frac{1}{100}(1 + \cdots + 100) = 50.5 \\ \operatorname{Var} X_i &= \frac{1}{100}(1^2 + \cdots + 100^2) - 50.5^2 = 833.25 \quad . \end{aligned}$$

Calculation of E X_i and Var X_i in R:

```
X6_27 <- 1:100
f_x6_27 <- rep(x = 1/100, times = 100)

Mean_X <- sum(f_x6_27 * X6_27)
Var_X <- sum(f_x6_27 * (X6_27 - Mean_X)^2)
```

Mean_X	Var_X
[1] 50.5	[1] 833.25

Notes

Example 6.27: see Ex. 6.26 b) (2)

Wirtschafts- und Sozialstatistik
Universität Trier

Now we are interested in the sample mean $\overline{X} = \frac{1}{3}(X_1 + X_2 + X_3)$. We get:

$$\mathrm{E}\,\overline{X} = \frac{1}{3}(\mathrm{E}\,X_1 + \mathrm{E}\,X_2 + \mathrm{E}\,X_3) = 50.5$$

$$\mathrm{Var}\,\overline{X} = \left(\frac{1}{3}\right)^2 (\mathrm{Var}\,X_1 + \mathrm{Var}\,X_2 + \mathrm{Var}\,X_3) = \frac{1}{3} \cdot 833.25 = 277.75 \quad .$$

Notice that the draws are stochastically independent (with replacement).
In the model without replacement we would have $\mathrm{E}\,\overline{X} = 50.5$ and $\mathrm{Var}\,\overline{X} = 272.139$.

Notes

Example 6.28: see Ex. 6.21 (1)

Wirtschafts- und Sozialstatistik
Universität Trier

We are now interested in $Y = 2X_1 + 4X_2 - 1$. We get:

$$\begin{aligned}\mathrm{E}\,Y &= 2\mathrm{E}\,X_1 + 4\mathrm{E}\,X_2 - 1\\ &= 2 \cdot 0.9 + 4 \cdot 0.6 - 1 = 3.2\end{aligned}$$

and

$$\begin{aligned}\mathrm{Var}\,Y &= \sum_{i=1}^{2}\sum_{j=1}^{2} a_i a_j \cdot \mathrm{Cov}\,(X_i, X_j)\\ &= 2^2 \cdot \mathrm{Var}\,X_1 + 2 \cdot 2 \cdot 4 \cdot \mathrm{Cov}\,(X_1, X_2) + 4^2 \cdot \mathrm{Var}\,X_2\\ &= 4 \cdot 0.63 - 16 \cdot 0.18 + 16 \cdot 0.48\\ &= 2.52 + 4.8 = 7.32 \quad .\end{aligned}$$

Notes

Example 6.28: see Ex. 6.21 (2)

Wirtschafts- und Sozialstatistik
Universität Trier

Calculation of E Y and Var Y in R

```
a0 <- -1
a1 <- 2
a2 <- 4

# ATTENTION: Means etc. from Ex. 6.21!
Mean_Y <- a0 + a1 * Mean_X1_old + a2 * Mean_X2_old

Var_Y <- a1^2 * Var_X1 + 2 * a1 * a2 * Cov_X1_X2_old +
         a2^2 * Var_X2
```

```
Mean_Y
[1] 3.2
```

```
Var_Y
[1] 7.32
```


Notes

Notes