Statistik

für Betriebswirtschaft, Internationales Management, Wirtschaftsinformatik und Informatik

Sommersemester 2016

Veranstaltungen zur Statistik für BW/IM Sommersemester 2016				
Was?	Wer?	Wann?	Wo?	ab wann?
Vorlesung Statistik	Etschberger	Mi, 14.00-17.00	B2.14	16.03.2016
Vorlesung Statistik PLUS	Etschberger/Jansen	- Blocktermin -	?	?
Übung Statistik	Etschberger	Mi, 17.00-18.30	A1.10	30.03.2016
Übung Statistik	Jansen	Di, 11.30-13.00	W1.06	22.03.2016
Übung Statistik	Jansen	Di, 14.00-15.30	W2.14	22.03.2016
Übung Statistik	Jansen	Mi, 11.30-13.00	W2.11	30.03.2016
Übung Statistik	Jansen	Do, 14.00-15.30	W2.14	31.03.2016
Übung Statistik	Schneller	Do, 14.00-15.30	W3.03	31.03.2016
Übung Statistik	Schneller	Do, 15.30-17.00	W3.03	31.03.2016
Übung Statistik	Wins	Di, 14.00-15.30	J3.19	22.03.2016
Übung Statistik	Wins	Di, 15.30-17.00	J3.19	22.03.2016
Offener Statistikraum	Jansen/Tutoren	Mo, 14.00-17.45	B3.05	04.04.2016
Veranstaltungen für Teilnehmer der WiMa-Klausur im Juli 2016				
Was?	Wer?	Wann?	Wo?	ab wann?
Tutorium Mathematik	Burkart	Do 13.30-15.00	W1.06	07.04.2016
Tutorium Mathematik	Burkart	Do 15.00-16.15	W1.06	07.04.2016
Offener Matheraum	Etschberger/Tutoren	Fr 11.30-15.30	B3.05	29.04.2016

Datum	Statistik für IM/BW	Nr.
Mittwoch, 16. März 2016	Einführung, R Installation, Rstudio Einführung, Skalen	1
Mittwoch, 23. März 2016	univ. deskr. Stat., Quantile, Plots	2
Mittwoch, 30. März 2016	Streuung, Konzentrationsmaße	3
Mittwoch, 6. April 2016	Kontingenztabellen, Mosaikplots, Korrelation	4
Mittwoch, 13. April 2016	Preisindizes, lineare Regression	5
Mittwoch, 20. April 2016	Kombinatorik, Wahrscheinlichkeit	6
Mittwoch, 27. April 2016	Wahrscheinlichkeit	7
Mittwoch, 4. Mai 2016	Pyramid	
Mittwoch, 11. Mai 2016	diskrete Zufallsvariablen	8
Mittwoch, 18. Mai 2016	Stetige ZV, Gleichverteilung	9
Mittwoch, 25. Mai 2016	Normalverteilung, Verteilungsparameter	10
Mittwoch, 1. Juni 2016	Schätzfunktionen und Punktschätzer, Konfidenzintervalle	11
Mittwoch, 8. Juni 2016	Konfidenzintervalle, Tests	12
Mittwoch, 15. Juni 2016	Go-Out-Tag	
Mittwoch, 22. Juni 2016	Tests, WH, (Fragen zur Probekl. In den Übungsgruppen)	13
Mittwoch, 29. Juni 2016	AW Prüfungswoche	

- Offener Stat-/Mathraum am 27.5. (Freitag) offen
- Nur zwei Anmeldungen für STAT PLUS --> Kurs findet nicht statt!
- HA 25.5.2016: 51, 66-71, 72d,e, 75-77

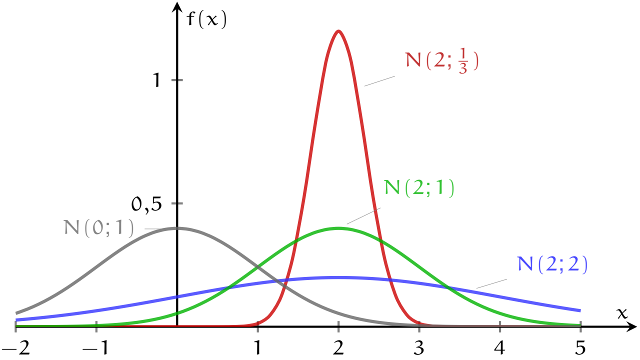
Prof. Dr. Stefan Etschberger
Hochschule Augsburg

Normalverteilung

Eine Zufallsvariable X mit einer Dichtefunktion

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$$

und $\sigma > 0$ heißt **normalverteilt**.



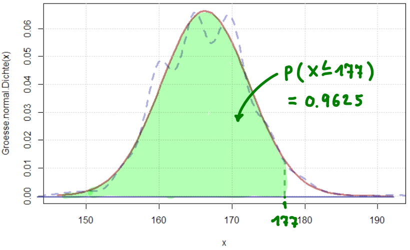
Kurzschreibweise: $X \sim N(\mu; \sigma)$

Beispiel (Größe der Mütter):

$X \sim N(\mu = 166.3\ ;\ \sigma = 6.02)$

$$P(X \leq 177) = F(177) = \Phi\left(\frac{177-\mu}{\sigma}\right)$$

$$\approx \Phi(1.78) = 0.9625$$



$$P(X \leq 152) = F(152) = \Phi\left(\frac{152-\mu}{\sigma}\right)$$

$$= \Phi(-2.38) = 1 - \Phi(2.38)$$

$$= 1 - 0.9913 = 0.0087$$

Die größten 10%: $P(X \geq x) = 0.1$

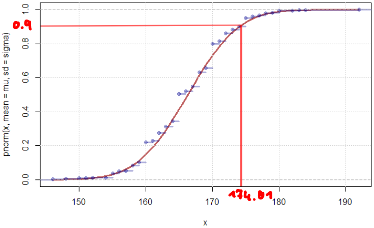
$$\Leftrightarrow 1 - P(X \leq x) = 0.1 \Leftrightarrow P(X \leq x) = 0.9$$

$$\Leftrightarrow F(x) = 0.9 \Leftrightarrow \Phi\left(\frac{x-\mu}{\sigma}\right) = 0.9 \Leftrightarrow \frac{x-\mu}{\sigma} = 1.28$$

$$\Leftrightarrow x = 1.28 \cdot \sigma + \mu = 1.28 \cdot 6.02 + 166.3 \approx 174.01$$

$x_1 \backslash x_2$	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6737	0.6773	0.6808	0.6844
0.5	0.6915	0.6950	0.6985	0.7020	0.7054	0.7089	0.7123	0.7157	0.7191
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518
0.7	0.7580	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823
0.8	0.7882	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8079	0.8106
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8290	0.8315	0.8340	0.8365
1	0.8414	0.8438	0.8461	0.8485	0.8508	0.8532	0.8554	0.8577	0.8599
1.1	0.8643	0.8665	0.8687	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997
1.3	0.9032	0.9049	0.9066	0.9083	0.9099	0.9115	0.9131	0.9147	0.9162
1.4	0.9193	0.9207	0.9222	0.9237	0.9251	0.9265	0.9279	0.9292	0.9306
1.5	0.9332	0.9345	0.9358	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430
1.6	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9516	0.9526	0.9535
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9600	0.9608	0.9616	0.9625
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9679	0.9686	0.9693	0.9700

$\Phi(1.78)$



```
# ------------------------------------------------------
# Statistik 25.5.2016
# ------------------------------------------------------

# Lade Daten
D <- read.csv("http://goo.gl/KCGF16", sep=";", dec=",")

# Groesse der Mütter
# Lösche fehlende Werte
Groesse = na.omit(D$GroesseM)

table(Groesse)

# korrigiere (wahrscheinlichen) Tippfehler
Groesse[Groesse < 100] = 100 + Groesse[Groesse < 100]

# Arithm. Mittel und (emp.) Standardabweichung als
# Schätzwert für mü und sigma
mu = mean(Groesse)
sigma = sqrt(mean((Groesse-mean(Groesse))^2))
data.frame(sigma, mu)

# Graph der Verteilungsfunktion (theoretisch und empirisch)
x = seq(from=min(Groesse), to=max(Groesse), length=100)
plot(x, pnorm(x, mean = mu, sd = sigma), type="l", lty=2, lwd=3, col = "#99000080")
grid()
Groesse.F = ecdf(Groesse)
plot(Groesse.F, add=TRUE, lwd=3, col="#000090A0")

# Linearisierter qq-plot
qqnorm(Groesse, pch=20, col="#00009050");
qqline(Groesse, col = "#99000090",lwd=3,lty=2)

Z = (177 - mu)/sigma
Z
pnorm(Z)
pnorm(177, mean = mu, sd=sigma)
# empirisch:
Groesse.F(177)

Z = (152 - mu)/sigma
Z
pnorm(Z)
pnorm(152, mean = mu, sd=sigma)
# empirisch:
Groesse.F(152)

# Quantile
qnorm(0.9, mean=mu, sd=sigma)

# 1-Sigma-Bereich
pnorm(mu + sigma, mean=mu, sd=sigma) - pnorm(mu - sigma, mean=mu, sd=sigma)
Groesse.F(mu + sigma) - Groesse.F(mu - sigma)

# 2-Sigma-Bereich
pnorm(mu + 2*sigma, mean=mu, sd=sigma) - pnorm(mu - 2*sigma, mean=mu, sd=sigma)
Groesse.F(mu + 2*sigma) - Groesse.F(mu - 2*sigma)

# 3-Sigma-Bereich
pnorm(mu + 3*sigma, mean=mu, sd=sigma) - pnorm(mu - 3*sigma, mean=mu, sd=sigma)
Groesse.F(mu + 3*sigma) - Groesse.F(mu - 3*sigma)
```

Normalverteilung: Gaußkurve

Normalverteilung

C.F. Gauß

Verteilungsfunktion Φ der Standardnormalverteilung

Dabei bedeutet $\Phi(x)$ zum Beispiel: $\Phi(2{,}13) = \Phi(2{,}1 + 0{,}03) = 0{,}9834$. Diesen Wert findet man in der Zeile mit $x_1 = 2{,}1$ und der Spalte mit $x_2 = 0{,}03$.

$x_1 \backslash x_2$	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5754
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6737	0.6773	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7020	0.7054	0.7089	0.7123	0.7157	0.7191	0.7224
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.7	0.7580	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7882	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8079	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8290	0.8315	0.8340	0.8365	0.8389
1	0.8414	0.8438	0.8461	0.8485	0.8508	0.8532	0.8554	0.8577	0.8599	0.8622
1.1	0.8643	0.8665	0.8687	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9083	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9193	0.9207	0.9222	0.9237	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9358	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9516	0.9526	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9600	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9679	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9865	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9914	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9933	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9975	0.9975	0.9976	0.9977	0.9978	0.9978	0.9979	0.9980	0.9980	0.9981
2.9	0.9981	0.9982	0.9983	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995	0.9995

Eigenschaften der Normalverteilung

- Dichte ist symmetrisch zu μ:

$$f(\mu - x) = f(\mu + x)$$

➡ μ ist Lage-, σ ist Streuungsparameter

- Standardnormalverteilung:
  $N(0;1)$ mit Verteilungsfunktion $\Phi(x)$ ($\rightarrow$ Tabelle 3)
- Kenntnis von $\Phi(x)$, μ und σ genügt, denn:
  $X \sim N(\mu;\sigma) \iff \frac{X-\mu}{\sigma} \sim N(0;1) \quad \Rightarrow$

$$\boxed{F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)}$$

- Tabelle enthält nur positive x: Deswegen

$$\boxed{\Phi(-x) = 1 - \Phi(x)}$$

Beispiel:

μ $\quad$ σ (Standard-abweichung)

Projektdauer $X \sim N(39; 2)$.

Wahrscheinlichkeit für Projektdauer zwischen 37 und 41 Wochen?

Lösung:

$$
\begin{aligned}
P(37 \leqq X \leqq 41) &= F(41) - F(37) \\
&= \Phi\left(\tfrac{41-39}{2}\right) - \Phi\left(\tfrac{37-39}{2}\right) \\
&= \Phi(1) - \Phi(-1) \\
&= \Phi(1) - [1 - \Phi(1)] \\
&= 2 \cdot \Phi(1) - 1 \\
&= 2 \cdot 0{,}8413 - 1 \\
&= 0{,}6826
\end{aligned}
$$

a) **Modus** x_{Mod}: $f(x_{\text{Mod}}) \geqq f(x)$ für alle x
(i.A. nicht eindeutig, z.B. Gleichverteilung)

Beispiele:

- Normalverteilung: $x_{\text{Mod}} = \mu$
- Diskrete Verteilung mit:

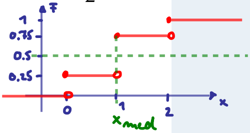
x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$\Rightarrow x_{\text{Mod}} = 1$

b) **Median** x_{Med}: $F(x_{\text{Med}}) = \frac{1}{2}$ bzw. kleinstes x mit $F(x) > \frac{1}{2}$

Beispiele:

- Normalverteilung: $x_{\text{Med}} = \mu$
- Diskrete Verteilung
oben: $F(0) = \frac{1}{4} < \frac{1}{2}$, $F(1) = \frac{3}{4} > \frac{1}{2} \Rightarrow x_{\text{Med}} = 1$



c) **α-Fraktil** x_α: $F(x_\alpha) = \alpha$ (für stetige Verteilungen)

Beispiel: $X \sim N(0;1)$, $Y \sim N(3;2)$

$$
\begin{aligned}
x_{0,975} &= & 1{,}96 \\
x_{0,025} &= -x_{0,975} &= -1{,}96 \\
y_{0,025} &= 2 \cdot x_{0,025} + 3 &= -0{,}92
\end{aligned}
$$

(Ta

$x_1 \backslash x_2$	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5754
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6737	0.6773	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7020	0.7054	0.7089	0.7123	0.7157	0.7191	0.7224
0.6	0.7258	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7518	0.7549
0.7	0.7580	0.7612	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7882	0.7910	0.7939	0.7967	0.7996	0.8023	0.8051	0.8079	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8290	0.8315	0.8340	0.8365	0.8389
1	0.8414	0.8438	0.8461	0.8485	0.8508	0.8532	0.8554	0.8577	0.8599	0.8622
1.1	0.8643	0.8665	0.8687	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9083	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9193	0.9207	0.9222	0.9237	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9358	0.9370	0.9382	0.9394	0.9406	0.9418	0.9430	0.9441
1.6	0.9452	0.9463	0.9474	0.9485	0.9495	0.9505	0.9516	0.9526	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9600	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9679	0.9686	0.9693	0.9700	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9762	0.9767
2	0.9773	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817

$\Phi(x_{0.975}) = 0.975$

$\Leftrightarrow$ $x_{0.975} = 1.96$

Hinweise:

- $x_{Med} = x_{0,5}$
- Wenn x_α nicht vertafelt $\rightarrow$ Interpolation:

$$
x_\alpha \approx x_a + (x_b - x_a) \cdot \frac{\alpha - a}{b - a}
$$

mit
a : größte vertafelte Zahl $< \alpha$
b : kleinste vertafelte Zahl $> \alpha$

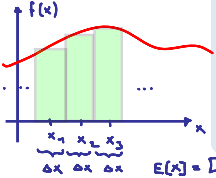
Beispiel: $X \sim N(0;1)$; $x_{0,6} \approx$
$0{,}25 + (0{,}26 - 0{,}25) \cdot \frac{0{,}6 - 0{,}5987}{0{,}6026 - 0{,}5987} = 0{,}2533$

Lageparameter: Erwartungswert

Ausprägung   rel. Häufigkeit

deskriptiven Statistik: $\bar{x} = \frac{1}{n}\sum x_i = \frac{1}{n}\sum a_j \cdot h_j = \sum a_j \cdot f_j$

d) Erwartungswert $E(X)$ bzw. μ:



$$E(X) = \begin{cases} \sum\limits_i x_i f(x_i), & \text{falls } X \text{ diskret} \\ \int\limits_{-\infty}^{\infty} x f(x)\,dx, & \text{falls } X \text{ stetig} \end{cases}$$

$E[x] = \sum x_i \cdot f(x_i) \cdot \Delta x$

Beispiel: Diskrete Verteilung mit

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$$\Rightarrow \quad E(X) = 0 \cdot \tfrac{1}{4} + 1 \cdot \tfrac{1}{2} + 2 \cdot \tfrac{1}{4} = 1$$

Beispiel: Für eine exponentialverteilte Zufallsvariable X mit der Dichte

$$f(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{für } x \geqslant 0 \\ 0 & \text{sonst} \end{cases} \quad \text{folgt}$$

$$E(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \lambda \int_0^{\infty} x \cdot e^{-\lambda x}\,dx = \lambda \left[-\frac{1}{\lambda} x e^{-\lambda x} - \int_0^{\infty} 1 \cdot \left(-\frac{1}{\lambda} e^{-\lambda x} \right) dx \right]$$

$$= -x e^{-\lambda x} - \frac{1}{\lambda} e^{-\lambda x} \Big|_0^{\infty} = -0 - \left(-0 - \frac{1}{\lambda} \right) = \frac{1}{\lambda}$$

Beispiel: Glücksspiel, Einsatz 1€

Auszahlungsverteilung ($X \hat{=}$ Gewinn)

x	-1	0	10
$f(x)$	0.9	0.08	0.02

Erwartungswert:

$$E[X] = -1 \cdot 0.9 + 0 \cdot 0.08 + 10 \cdot 0.02$$
$$= -0.70 €$$

Beispiel: Erwartungswert der Gleichverteilung

X gleichverteilt mit $f(x) = \begin{cases} \frac{1}{b-a} & \text{f. } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)\,dx = \int_a^b x \cdot \left(\frac{1}{b-a}\right) dx$$

$$= \frac{1}{b-a} \int_a^b x\,dx = \frac{1}{b-a} \cdot \left[\tfrac{1}{2} x^2\right]_a^b$$

$$= \frac{1}{b-a} \cdot \tfrac{1}{2} \cdot (b^2 - a^2)$$

$$= \frac{b^2 - a^2}{(b-a) \cdot 2} = \frac{\cancel{(b-a)}(b+a)}{\cancel{(b-a)} \cdot 2} = \frac{b+a}{2}$$

$$\Rightarrow \boxed{E[X] = \frac{b+a}{2}}$$

Rechenregeln für den Erwartungswert

- Ist f symmetrisch bzgl. a, so gilt $E(X) = a$
  Beispiel: f der Gleichverteilung symmetrisch bzgl. $\frac{a+b}{2} \Rightarrow E(X) = \frac{a+b}{2}$
- Lineare Transformation:

$$\boxed{E(a + bX) = a + b \cdot E(X)}$$

- Summenbildung:

$$\boxed{E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)}$$

Beispiel: X gleichverteilt in $[0; 10]$, $Y \sim N(1; 1)$; $Z = X + 5Y$

$$E(Z) = E(X + 5Y) = E(X) + E(5Y) = E(X) + 5 \cdot E(Y) = \frac{10+0}{2} + 5 \cdot 1 = 10$$

- Unabhängigkeit:

$$X,\ Y \text{ unabhängig} \Rightarrow E(X \cdot Y) = E(X) \cdot E(Y)$$

Streuungsparameter

deskriptiv: mittlere quadratische Abw. $\frac{1}{n}\sum(x_i-\bar{x})^2$

$= \frac{1}{n}\sum(a_j-\bar{x})^2\cdot h_j$

$= \sum(a_j-\bar{x})^2\cdot f_j$

- **Varianz** $\text{Var}(X)$ bzw. σ^2:

$$\text{Var}(X) = E([X-E(X)]^2) = \begin{cases} \sum\limits_i [x_i - E(X)]^2 f(x_i), & \text{wenn } X \text{ diskret} \\ \int_{-\infty}^{\infty} [x-E(X)]^2 f(x)\,dx, & \text{wenn } X \text{ stetig} \end{cases}$$

- **Standardabweichung** $\text{Sta}(X)$ bzw. σ: $\text{Sta}(X) = \sqrt{\text{Var}(X)}$

Beispiel: Diskrete Verteilung

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

:

$$\text{Var}(X) = (0-1)^2\cdot\frac{1}{4} + (1-1)^2\cdot\frac{1}{2} + (2-1)^2\cdot\frac{1}{4} = \frac{1}{2}$$

Beispiel: Für eine exponentialverteilte Zufallsvariable X (Dichte siehe Erwartungswert) folgt

$$\text{Var}(X) = \int_{-\infty}^{\infty} (x-E(X))f(x)\,dx = \lambda\int_0^{\infty}\left(x-\tfrac{1}{\lambda}\right)^2\cdot e^{-\lambda x}\,dx$$

$$= e^{-\lambda x}\left(-x^2 + \tfrac{2x}{\lambda} - \left(\tfrac{1}{\lambda}\right)^2 - \tfrac{2}{\lambda^2} - \tfrac{2x}{\lambda} + \tfrac{2}{\lambda^2}\right)\Big|_0^{\infty}$$

$$= 0 - \left(-0^2 - \left(\tfrac{1}{\lambda}\right)^2\right) = \frac{1}{\lambda^2}$$

Beispiel: 1-/2-/3-Sigma-Bereich der Normalverteilung

$X \sim N(\mu = 166{,}3 \; ; \; \sigma = 6{,}02)$

$P(\mu - \sigma \leq X \leq \mu + \sigma)$

$= P(160{,}28 \leq X \leq 172{,}32)$

$= F(\mu + \sigma) - F(\mu - \sigma)$

$= \Phi\left(\frac{\mu + \sigma - \mu}{\sigma}\right) - \Phi\left(\frac{\mu - \sigma - \mu}{\sigma}\right)$

$= \Phi(1) - \Phi(-1) = \Phi(1) - (1 - \Phi(1))$

$= 2 \cdot \Phi(1) - 1 = 2 \cdot 0.84134 - 1 = 0.68264$

(W. für 1-Sigma Bereich)

$P(\mu - 2\sigma \leq X \leq \mu + 2\sigma)$

$= 2 \cdot \Phi(2) - 1 = 2 \cdot 0.97725 - 1 = 0.9545$

W. für 2-Sigma-Bereich

$P(\mu - 3\sigma \leq X \leq \mu + 3\sigma)$

$= 2 \cdot \Phi(3) - 1 = 2 \cdot 0.99865 - 1 = 0.9973$

W. für 3-Sigma-Bereich

Beispiel: Varianz der Gleichverteilung

X mit $f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$

$Var[X] = E[X^2] - E^2[X]$

$= \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx - \left(\frac{a+b}{2}\right)^2$

$= \int_a^b x^2 \cdot \left(\frac{1}{b-a}\right) dx - \ldots$

$= \frac{1}{b-a} \left[\frac{1}{3} x^3\right]_a^b - \ldots$

$= \frac{1}{b-a} \cdot \frac{1}{3} \cdot (b^3 - a^3) - \frac{1}{4} \cdot (a+b)^2$

$= \frac{1}{3} \cdot (b^2 + ab + a^2) - \frac{1}{4} (a^2 + 2ab + b^2)$

$= \frac{1}{12} \cdot (4b^2 + 4ab + 4a^2 - 3a^2 - 6ab - 3b^2)$

$= \frac{1}{12} \cdot (b^2 - 2ab + a^2) = \frac{(b-a)^2}{12}$

$Var[X] = \frac{(b-a)^2}{12}$

$$
\begin{array}{l}
(b^3 - a^3) : (b - a) = b^2 + ab + a^2 \\
\underline{-(b^3 - ab^2)} \\
\quad ab^2 - a^3 \\
\quad \underline{-(ab^2 - a^2b)} \\
\qquad a^2b - a^3 \\
\qquad \underline{-(a^2b - a^3)} \\
\qquad\quad 0
\end{array}
$$

Rechenregeln für die Varianz

- **Verschiebungssatz**:

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - [\mathrm{E}(X)]^2$$

Beispiel: Diskrete Verteilung

x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

:

$$\begin{aligned} \mathrm{E}(X^2) &= 0^2 \cdot \frac{1}{4} + 1^2 \cdot \frac{1}{2} + 2^2 \cdot \frac{1}{4} \\ &= \frac{3}{2} \\ \Rightarrow \quad \mathrm{E}(X^2) - [\mathrm{E}(X)]^2 &= \frac{3}{2} - 1^2 = \frac{1}{2} = \mathrm{Var}(X) \end{aligned}$$

- **Lineare Transformation**:

$$\mathrm{Var}(a + bX) = b^2\, \mathrm{Var}(X)$$

- **Summenbildung** gilt nur, wenn die X_i unabhängig! Dann:

$$\mathrm{Var}\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} \mathrm{Var}(X_i)$$

Erwartungswerte und Varianzen wichtiger Verteilungen

Verteilung von X	$E(X)$	$Var(X)$
Binomialverteilung $B(n;p)$	np	$np(1-p)$
Hypergeometrische Verteilung mit den Parametern N, M, n	$n\frac{M}{N}$	$n\frac{M}{N}\frac{N-M}{N}\frac{N-n}{N-1}$
Poisson-Verteilung $P(\lambda)$	λ	λ
Gleichverteilung in $[a;b]$ mit $a < b$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Normalverteilung $N(\mu;\sigma)$	μ	σ^2

- Für beliebige Zufallsvariablen X und $\varepsilon > 0$ gilt die Ungleichung von Tschebyschow:

$$P\big(|X - E[X]| \geqslant \varepsilon\big) \leqslant \frac{\operatorname{Var}[X]}{\varepsilon^2}$$

Beispiele:

- X ist gleichverteilt mit Parametern a, b und $\varepsilon = \frac{1}{3}(a-b)$, also $E[X] = \frac{1}{2}(a+b)$ und $\operatorname{Var}[X] = \frac{1}{12}(a-b)^2$

$$\Rightarrow P\Big(\big|X - \tfrac{1}{2}(a+b)\big| \geqslant \tfrac{1}{3}(a-b)\Big) \leqslant \frac{(a-b)^2}{12} \cdot \frac{3^2}{(a-b)^2} = 3/4$$

- $X \sim B(100; 0{,}2)$ und $\varepsilon = 10$
damit: $E[X] = 100 \cdot 0{,}2 = 20$ und $\operatorname{Var}[X] = 100 \cdot 0{,}2 \cdot (1 - 0{,}2) = 16$

$$\Rightarrow P\big(|X - 20| \geqslant 10\big) \leqslant \frac{16}{10^2} = 0{,}16$$

Kovarianz und Korrelation

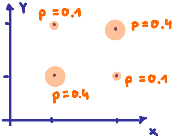
$$\frac{\sum (x_i-\bar{x})(y_i-\bar{y})}{\sqrt{\sum (x_i-\bar{x})^2 \sum (y_i-\bar{y})^2}}$$

- **Kovarianz:**

$$\begin{aligned} \operatorname{Cov}(X,Y) &= E[(X-E(X))(Y-E(Y))] \\ &= E(X\cdot Y) - E(X)\cdot E(Y) \\ &\text{(Verschiebungssatz)} \end{aligned}$$

- **Korrelationskoeffizient:**

$$\rho(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\cdot \operatorname{Var}(Y)}}$$



- **Bemerkungen:**
  - ρ ist r nachgebildet $\Rightarrow \rho \in [-1;1]$
  - $|\rho| = 1 \iff Y = a + bX$ (mit $b \neq 0$)
  - $\rho = 0 \iff X,\ Y$ **unkorreliert**
- Varianz einer Summe zweier ZV:

$$\operatorname{Var}(X+Y) = \operatorname{Var}(X) + \operatorname{Var}(Y) + 2\operatorname{Cov}(X,Y)$$

Beispiel

$Y \backslash X$	-1	0	2	
1	0.5	0	0.05	0.55
2	0	0.1	0.35	0.45
	0.5	0.1	0.4	

$$\mathrm{Cov}[X,Y] = E[X \cdot Y] - E[X] \cdot E[Y]$$

$$E[X] = -1 \cdot 0.5 + 0 \cdot 0.1 + 2 \cdot 0.4 = 0.3$$

$$E[Y] = 1 \cdot 0.55 + 2 \cdot 0.45 = 1.45$$

$$E[X \cdot Y] = (-1) \cdot 1 \cdot 0.5 + 0 \cdot 1 \cdot 0 + 2 \cdot 1 \cdot 0.05 + (-1) \cdot 2 \cdot 0 + 0 \cdot 2 \cdot 0.1 + 2 \cdot 2 \cdot 0.35$$

$$= 1.0$$

$$\mathrm{Cov}[X,Y] = 1.0 - 0.3 \cdot 1.45 = +0.565$$

$$\rho[X,Y] = \frac{\mathrm{Cov}[X,Y]}{\mathrm{Std}[X] \cdot \mathrm{Std}[Y]}$$

Verteilungsfunktion Φ der Standardnormalverteilung

Dabei bedeutet $\Phi(x)$ zum Beispiel: $\Phi(2{,}13) = \Phi(2{,}1 + 0{,}03) = 0{,}9834$. Diesen Wert findet man in der Zeile mit $x_1 = 2{,}1$ und der Spalte mit $x_2 = 0{,}03$.

$x_1 \backslash x_2$	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.50000	0.50399	0.50798	0.51197	0.51595	0.51994	0.52392	0.52790	0.53188	0.53586
0.1	0.53983	0.54380	0.54776	0.55172	0.55567	0.55962	0.56356	0.56749	0.57142	0.57535
0.2	0.57926	0.58317	0.58706	0.59095	0.59483	0.59871	0.60257	0.60642	0.61026	0.61409
0.3	0.61791	0.62172	0.62552	0.62930	0.63307	0.63683	0.64058	0.64431	0.64803	0.65173
0.4	0.65542	0.65910	0.66276	0.66640	0.67003	0.67364	0.67724	0.68082	0.68439	0.68793
0.5	0.69146	0.69497	0.69847	0.70194	0.70540	0.70884	0.71226	0.71566	0.71904	0.72240
0.6	0.72575	0.72907	0.73237	0.73565	0.73891	0.74215	0.74537	0.74857	0.75175	0.75490
0.7	0.75804	0.76115	0.76424	0.76730	0.77035	0.77337	0.77637	0.77935	0.78230	0.78524
0.8	0.78814	0.79103	0.79389	0.79673	0.79955	0.80234	0.80511	0.80785	0.81057	0.81327
0.9	0.81594	0.81859	0.82121	0.82381	0.82639	0.82894	0.83147	0.83398	0.83646	0.83891
1	0.84134	0.84375	0.84614	0.84850	0.85083	0.85314	0.85543	0.85769	0.85993	0.86214
1.1	0.86433	0.86650	0.86864	0.87076	0.87286	0.87493	0.87698	0.87900	0.88100	0.88298
1.2	0.88493	0.88686	0.88877	0.89065	0.89251	0.89435	0.89617	0.89796	0.89973	0.90147
1.3	0.90320	0.90490	0.90658	0.90824	0.90988	0.91149	0.91309	0.91466	0.91621	0.91774
1.4	0.91924	0.92073	0.92220	0.92364	0.92507	0.92647	0.92785	0.92922	0.93056	0.93189
1.5	0.93319	0.93448	0.93574	0.93699	0.93822	0.93943	0.94062	0.94179	0.94295	0.94408
1.6	0.94520	0.94630	0.94738	0.94845	0.94950	0.95053	0.95154	0.95254	0.95352	0.95449
1.7	0.95543	0.95637	0.95728	0.95818	0.95907	0.95994	0.96080	0.96164	0.96246	0.96327
1.8	0.96407	0.96485	0.96562	0.96638	0.96712	0.96784	0.96856	0.96926	0.96995	0.97062
1.9	0.97128	0.97193	0.97257	0.97320	0.97381	0.97441	0.97500	0.97558	0.97615	0.97670
2	0.97725	0.97778	0.97831	0.97882	0.97932	0.97982	0.98030	0.98077	0.98124	0.98169
2.1	0.98214	0.98257	0.98300	0.98341	0.98382	0.98422	0.98461	0.98500	0.98537	0.98574
2.2	0.98610	0.98645	0.98679	0.98713	0.98745	0.98778	0.98809	0.98840	0.98870	0.98899
2.3	0.98928	0.98956	0.98983	0.99010	0.99036	0.99061	0.99086	0.99111	0.99134	0.99158
2.4	0.99180	0.99202	0.99224	0.99245	0.99266	0.99286	0.99305	0.99324	0.99343	0.99361
2.5	0.99379	0.99396	0.99413	0.99430	0.99446	0.99461	0.99477	0.99492	0.99506	0.99520
2.6	0.99534	0.99547	0.99560	0.99573	0.99585	0.99598	0.99609	0.99621	0.99632	0.99643
2.7	0.99653	0.99664	0.99674	0.99683	0.99693	0.99702	0.99711	0.99720	0.99728	0.99736
2.8	0.99744	0.99752	0.99760	0.99767	0.99774	0.99781	0.99788	0.99795	0.99801	0.99807
2.9	0.99813	0.99819	0.99825	0.99831	0.99836	0.99841	0.99846	0.99851	0.99856	0.99861
3	0.99865	0.99869	0.99874	0.99878	0.99882	0.99886	0.99889	0.99893	0.99897	0.99900
3.1	0.99903	0.99906	0.99910	0.99913	0.99916	0.99918	0.99921	0.99924	0.99926	0.99929
3.2	0.99931	0.99934	0.99936	0.99938	0.99940	0.99942	0.99944	0.99946	0.99948	0.99950
3.3	0.99952	0.99953	0.99955	0.99957	0.99958	0.99960	0.99961	0.99962	0.99964	0.99965
3.4	0.99966	0.99968	0.99969	0.99970	0.99971	0.99972	0.99973	0.99974	0.99975	0.99976