Anmerkungen zur Vorlesung Statistik vom 06.12.2016

Inhalt:

- Stetige Zufallsvariablen
  - Normalverteilung (Eigenschaften, R)
- Verteilungsparameter:
  - Lageparameter
    - Modus
    - Median
    - α-Fraktil
    - Erwartungswert und Rechenregeln
  - Streuungsparameter
    - Varianz und Rechenregeln
- Kovarianz, Korrelation und Unabhängigkeit von Zufallsvariablen

Aufgabensammlung Statistik:

06.12.2016
Hausaufgabe
A61-A65 sowie
A66-A77

Klausurtermin: Freitag, 13. Januar 2017

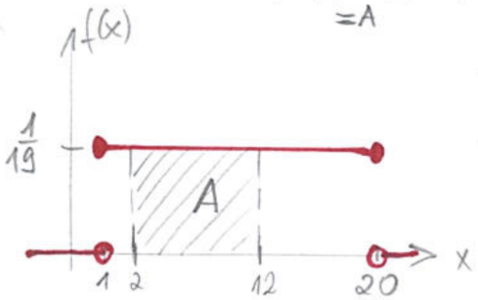
Gleichverteilung

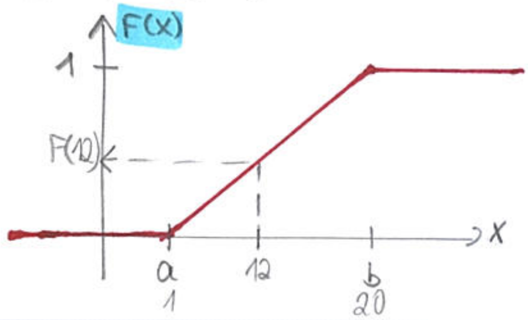
▶ Verteilungsfunktion der Gleichverteilung:

$$F(x) = \begin{cases} 0 \quad , & \text{falls} \quad x < a \\ \dfrac{x-a}{b-a}, & \text{falls} \quad a \leqq x \leqq b \\ 1 \quad , & \text{falls} \quad x > b \end{cases}$$

▶ **Beispiel:** X gleichverteilt in $[1;20]$

$$P(\underbrace{2 \leqq X \leqq 12}_{=A}) = F(12) - F(2) = \frac{12-1}{20-1} - \frac{2-1}{20-1}$$
$$= \frac{12-2}{20-1} = \frac{10}{19}$$
$$= 0{,}5263$$



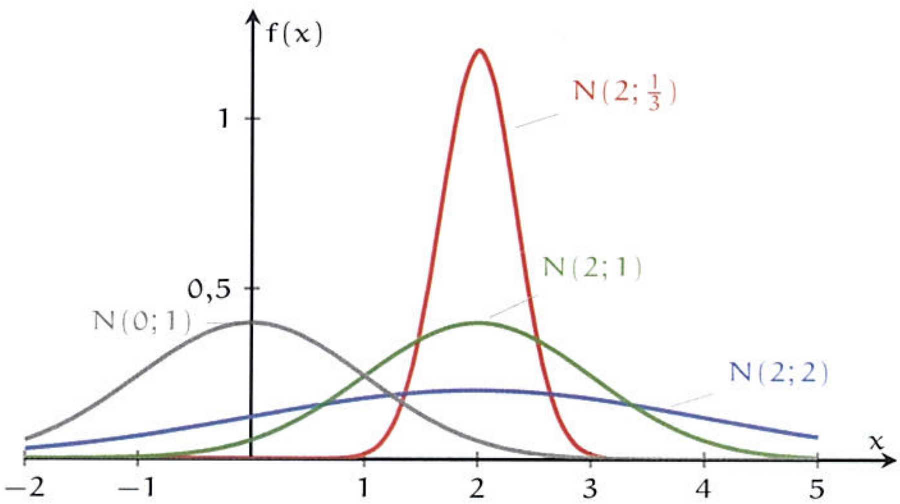


Normalverteilung

Eine Zufallsvariable X mit einer Dichtefunktion

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

und $\sigma > 0$ heißt normalverteilt.



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

Lageparameter

Streuungsparameter

Normalverteilung: Gaußkurve

Carl Friedrich

Normalverteilung

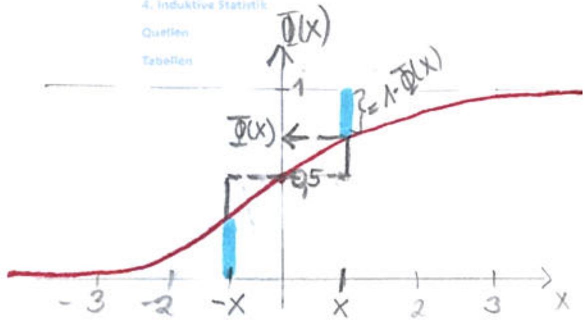
C.F. Gauß

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 $x_2 = 0{,}03$.

$\Phi(1{,}01) =$

$x_1 \setminus x_2$	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) F.218
- Kenntnis von $\Phi(x)$, μ und σ genügt, denn:
  $X \sim N(\mu; \sigma) \iff \frac{X-\mu}{\sigma} \sim N(0; 1) \quad \Rightarrow$

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

- Tabelle enthält nur positive x: Deswegen

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

Es gilt auch:
i) $F(x) = 0.5 \rightarrow x = \mu$
ii) $F(x_1) = 1 - F(x_2) \rightarrow \mu = \frac{x_1 + x_2}{2}$ $(x_1 \leq x_2)$

Normalverteilung: Beispiel

Beispiel:

Projektdauer $X \sim N(\overset{\mu}{39}; \overset{\sigma}{2})$.

Wahrscheinlichkeit für Projektdauer zwischen 37 und 41 Wochen?

Lösung:

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

weitere mögliche Fragestellungen:
i) $P(X \leq 40)$ (höchstens 40 Wochen / weniger als 40 Wochen)
ii) $P(X \geq x) \geq 0.10$ (Die "höchsten" 10%)
...

$X \sim N(39, 2)$

$P(X \geq x) \geq 0.10 \Leftrightarrow 1 - P(X \leq x) \geq 0.10$

$\Leftrightarrow P(X \leq x) \leq 0.90$

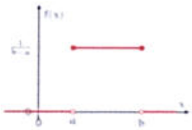
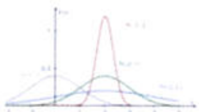
$\Leftrightarrow F(x) = \Phi\left(\frac{x-39}{2}\right) \leq 0.90$

Tabelle

$\Phi(1{,}28) = 0.8997$ $\Leftrightarrow \frac{x-39}{2} \leq 1{,}28$

$\Leftrightarrow X \leq 1{,}28 \cdot 2 + 39 = 41{,}56$

Stetige Zufallsvariablen/Verteilungen:

	Dichtefunktion	Verteilungsfunktion	E	Var
stetige Gleich-verteilung U(a; b)	$f(x) = \begin{cases} \frac{1}{b-a}, & \text{falls } a \leqq x \leqq b \\ 0, & \text{sonst} \end{cases}$ 	$F(x) = \begin{cases} 0, & \text{falls } x < a \\ \frac{x-a}{b-a}, & \text{falls } a \leqq x \leqq b \\ 1, & \text{falls } x > b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
	R dunif(x, min=0, max=1)	R punif(x, min=0, max=1)		
Normal-verteilung $N(\mu; \sigma)$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{(x-\mu)^2}{2\sigma^2}}$ 	$F(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$ ***Standardnormalverteilung N(0;1) ist tabelliert***	μ	σ^2
	Symmetrieeigenschaften: 1.) $\Phi(-x) = 1 - \Phi(x)$ 2.) $F(x) = 0.5 \rightarrow x = \mu$ 3.) $F(x_1) = 1 - F(x_2) \rightarrow \mu = \frac{x_1+x_2}{2}$			
	R dnorm(x, mean=0, sd=1)	R pnorm(x, mean=0, sd=1)		

```
#####
# 06.12.2016: metrische Zufalssvariablen - Normalverteilung
#####

mu=39
sigma=2
ZZ=rnorm(n=400, mean=mu, sd=sigma)
ZZ2=rnorm(n=100000, mean=mu, sd=sigma)

###empirische Verteilungsparameter:
#MW=mean(ZZ)
#STD=sqrt((length(ZZ)-1)/(length(ZZ))*var(ZZ))
#MW2=mean(ZZ2)
#STD2=sqrt((length(ZZ2)-1)/(length(ZZ2))*var(ZZ2))

###Dichtefunktion:
x=seq(29,49,by=0.05)
plot(x, dnorm(x, mean=mu, sd=sigma), main="Dichte", type="l", lwd=2, ylim=c(0,
0.25))  #theoretisch
lines(density(ZZ, bw=0.25), col=4, lwd=2)
      #empirisch
#lines(density(ZZ, bw=1), col=4, lwd=2)
      #empirisch
lines(density(ZZ2, bw=0.25), col=2, lwd=2)
      #empirisch

###Verteilungsfunktion:
plot(x, pnorm(x, mean=mu, sd=sigma), main="Verteilungsfunktion", type="l",
lwd=2)          #theoretisch
plot(ecdf(ZZ), col=4, add=T, lwd=2)
    #empirisch
plot(ecdf(ZZ2), col=2, add=T, lwd=2)
    #empirisch

###Linearisierter QQ-Plot
par(mfrow=c(2,1))
qqnorm(ZZ, pch=20)
qqline(ZZ, col=2, lwd=3, lty=2)

qqnorm(ZZ2, pch=20)
qqline(ZZ2, col=2, lwd=3, lty=2)


###k-Sigma Bereiche:
pnorm(mu+sigma, mean=mu, sd=sigma) - pnorm(mu-sigma, mean=mu, sd=sigma)
pnorm(mu+2*sigma, mean=mu, sd=sigma) - pnorm(mu-2*sigma, mean=mu, sd=sigma)
pnorm(mu+3*sigma, mean=mu, sd=sigma) - pnorm(mu-3*sigma, mean=mu, sd=sigma)
```

Lageparameter

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

Beispiele:

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

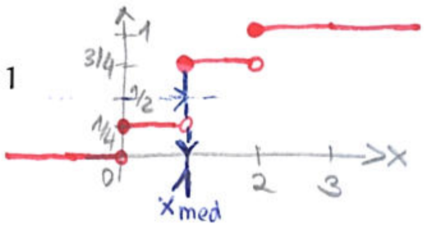
x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$\Rightarrow x_{Mod} = 1$

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

Beispiele:

- Normalverteilung: $x_{Med} = \mu$ $\Leftarrow F(\mu) = 0.5$
- Diskrete Verteilung
  oben: $F(0) = \frac{1}{4} < \frac{1}{2}$, $F(1) = \frac{3}{4} > \frac{1}{2} \Rightarrow x_{Med} = 1$

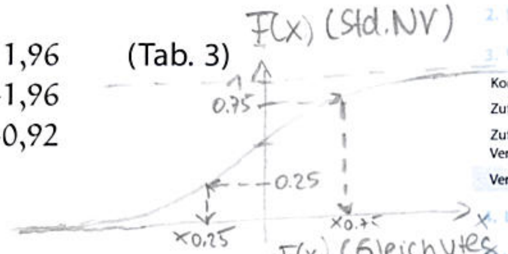


Lageparameter: Fraktile

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

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

$$
\begin{aligned}
x_{0,975} &= && 1{,}96 \quad \text{(Tab. 3)}\\
x_{0,025} &= -x_{0,975} &&= -1{,}96\\
y_{0,025} &= 2 \cdot x_{0,025} + 3 &&= -0{,}92
\end{aligned}
$$



Hinweise:

- $x_{Med} = x_{0,5}$
- Wenn x_α nicht vertafelt → 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: $X \sim N(0;1)$; $x_{0,6} \approx$
$0{,}25 + (0{,}26 - 0{,}25) \cdot \frac{0{,}6 - 0{,}5987}{0{,}6026 - 0{,}5987} = 0{,}2533$ (vgl. F218)

Lageparameter: Erwartungswert

aus deskriptiver Statistik: $\bar{x} = \frac{1}{n} \cdot \sum_i x_i = \frac{1}{n} \cdot \sum_j a_j \cdot h(a_j) = \sum_j a_j f(a_j)$

d) Erwartungswert $E(X)$ bzw. μ:

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

Beispiel: Diskrete Verteilung mit

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

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

Beispiel: Für eine exponentialverteilte Zufallsvariable X mit der Dichte *2

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

partielle Integration

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

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

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:

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

- Summenbildung:

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

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:

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

*¹ Beispiel: X diskret, E(X)?

Glücksspiel, Einsatz 1 €

Auszahlungsverteilung (X ≙ Gewinn)

x	-1	3	14
$f(x)$	$30/36$	$5/36$	$1/36$

$$\Rightarrow E(X) = -1 \cdot 30/36 + 3 \cdot 5/36 + 14 \cdot 1/36 = \underline{\underline{-1/36}}$$

*² Beispiel: X stetig, E(X)?

$X \sim U[a;b]$, d.h. $f(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \leq x \leq b \\ 0 & \text{sonst} \end{cases}$

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

$$= \frac{1}{b-a} \cdot \left[\frac{1}{2} x^2\right]_a^b = \frac{1}{b-a} \cdot \frac{1}{2} \cdot \left(b^2 - a^2\right) =$$

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

Streuungsparameter

aus deskriptiver Statistik: $MQA = s^2 = \frac{1}{n} \cdot \Sigma (x_i - \bar{x})^2 =$

- Varianz $\mathrm{Var}(X)$ bzw. σ^2: $= \frac{1}{n} \cdot \Sigma (a_j - \bar{x})^2 \cdot h(a_j) = \Sigma (a_j - \bar{x})^2 \cdot f(a_j)$

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

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

Beispiel: Diskrete Verteilung

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

:

$*^1$

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

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

$*^2$

$$\begin{aligned} \mathrm{Var}(X) &= \int_{-\infty}^{\infty} (x - E(X))^2 f(x)\,dx = \lambda \int_0^{\infty} \left(x - \tfrac{1}{\lambda}\right)^2 \cdot e^{-\lambda x}\,dx \\ &= e^{-\lambda x} \left(-x^2 + \tfrac{2x}{\lambda} - \left(\tfrac{1}{\lambda}\right)^2 - \tfrac{2}{\lambda^2} - \tfrac{2x}{\lambda} + \tfrac{2}{\lambda^2}\right)\Big|_0^{\infty} \\ &= 0 - \left(-0^2 - \left(\tfrac{1}{\lambda}\right)^2\right) = \frac{1}{\lambda^2} \end{aligned}$$

Rechenregeln für die Varianz

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

- Verschiebungssatz:

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

Beispiel: Diskrete Verteilung

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

:

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

- Lineare Transformation:

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

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

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

*1 Beispiel

$$Var(X) = \left(-1+\frac{1}{36}\right)^2 \cdot \frac{30}{36} + \left(3+\frac{1}{36}\right)^2 \cdot \frac{5}{36} + \left(14+\frac{1}{36}\right)^2 \cdot \frac{1}{36}$$

$$\approx 7{,}5270$$

*2 Beispiel:

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

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

$$= \left(\frac{1}{b-a}\right) \cdot \left[\frac{1}{3}x^3 - \left(\frac{b+a}{2}\right) x^2 + \left(\frac{b+a}{2}\right)^2 \cdot x \right]_a^b =$$

$$= \left(\frac{1}{b-a}\right) \left[\frac{1}{3}(b^3-a^3) - \left(\frac{b+a}{2}\right)(b^2-a^2) + \left(\frac{b+a}{2}\right)^2 (b-a) \right]$$

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

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

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

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

NR: $(b^3 - a^3) : (b-a) = b^2 + ab + a^2$

$-(b^3 - ab^2)$

$ab^2 - a^3$

$-(ab^2 - a^2b)$

$a^2b - a^3$

$-(a^2b - a^3)$

0

Mit Verschiebungssatz:

$$Var(X) = E(X^2) - (E(X))^2$$

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

$$= \left(\frac{1}{b-a}\right) \cdot \left[\frac{1}{3} x^3\right]_a^b - \left(\frac{a+b}{2}\right)^2 =$$

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

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

$$= \ldots = \frac{(b-a)^2}{12}$$

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 $a<b$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normalverteilung $N(\mu;\sigma)$	μ	σ^2

Anwendung: Ungleichung von Tschebyschow

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

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

Beispiele:

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

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

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

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

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	

$$Cov(X,Y) = E(X \cdot Y) - E(X) \cdot E(Y)$$

$$E(X) = \sum_i x_i \cdot f(x_i) = -1 \cdot 0.5 + 0 \cdot 0.1 + 2 \cdot 0.4 = 0.3$$

$$E(Y) = \sum_j y_j \cdot f(y_j) = 1 \cdot 0.55 + 2 \cdot 0.45 = 1.45$$

$$E(X \cdot Y) = \sum_{i,j} x_i \, y_j \, f(x_i, y_j) = -1 \cdot 1 \cdot 0.5 + 0 \cdot 1 \cdot 0 + 2 \cdot 1 \cdot 0.05$$

$$-1 \cdot 2 \cdot 0 + 0 \cdot 2 \cdot 0.1 + 2 \cdot 2 \cdot 0.35 = 1$$

$$\Rightarrow Cov(X,Y) = 1 - 0.3 \cdot 1.45 = 0.565$$

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sqrt{Var(X) \cdot Var(Y)}} = \frac{0.565}{\sqrt{2.1 \cdot 0.2475}} \simeq 0.7837$$

$$Var(X) = E(X^2) - (E(X))^2 = 2.1 - 0.3^2 = 2.01$$

$$E(X^2) = \sum_i x_i^2 \cdot f(x_i) = (-1)^2 \cdot 0.5 + 0^2 \cdot 0.1 + 2^2 \cdot 0.4 = 2.1$$

$$Var(Y) = E(Y^2) - (E(Y))^2 = 2.35 - 1.45^2 = 0.2475$$

$$E(Y^2) = \sum_j y_j^2 \cdot f(y_j) = 1^2 \cdot 0.55 + 2^2 \cdot 0.45 = 2.35$$

Kovarianz und Korrelation

- **Kovarianz:**

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

(Verschiebungssatz)

unter Unabhängigkeit: $\mathrm{Cov}(X,Y) = 0$ (vgl. F.154)

- **Korrelationskoeffizient:**

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

- **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:

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

unter Unabhängigkeit:
$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

Statistik: Table of Contents