7.2 NORMALVERTEILUNG

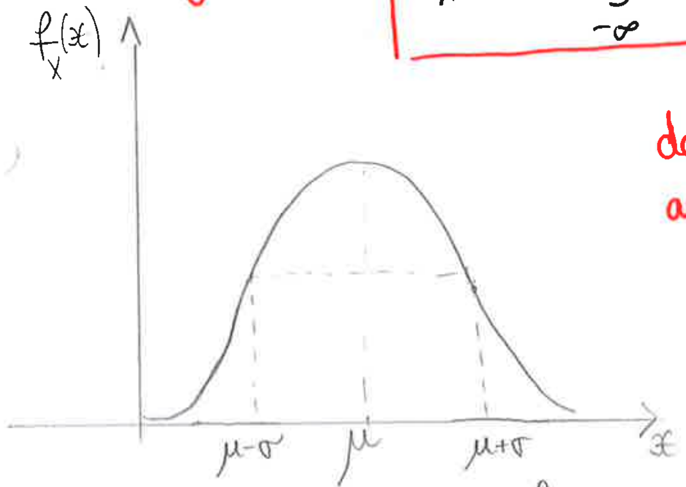
- Spielt in der W'theorie und Statistik eine zentrale Rolle.
- Wurde von Gauss für die Modellierung der Verteilung von Meßfehlern vorgeschlagen.
- Wird auch Gaußsche Verteilung genannt.
- Normalverteilung ist Grenzverteilung von Binomialverteilung.

Def: Eine Z.V $X:\Omega\to\mathbb{R}$ heißt Normalverteilt mit Parameter $\mu\in\mathbb{R}$ (Lokationsparameter) und $\sigma>0$ (Skalierungsparameter), falls X die folgende Dichtefkt. besitzt:

$$\boxed{f_X(x)=\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},\quad \text{für } x\in\mathbb{R},\ \mu\in\mathbb{R},\ \sigma>0}$$

Schreibweise: $X\sim\mathcal{N}(\mu,\sigma)$

Verteilungsfkt:
$$\boxed{F_X(x)=\int_{-\infty}^{x}\frac{1}{\sigma\sqrt{2\pi}}\,e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}dt}$$
→ lässt sich nur in Integralform angeben, da die Dichtefkt. nicht elementar ausintegrierbar ist.



Gaußsche Glockenkurve $f_X(x)$

Bemerkungen: Die Parameter μ und σ sind die Kennwerte von $X\sim\mathcal{N}(\mu,\sigma)$, d.h

$$\left\{\begin{array}{l} E[X]=\mu \\ Var(X)=\sigma^2 \end{array}\right.\quad ;\ \sigma=\sqrt{Var X}\ \Bigg\}\ \text{wird später bewiesen}$$

Folie zur Vorlesung
“Wahrscheinlichkeitsrechnung und Stoch. Prozesse”

Zur Normalverteilung:

Es gilt für jedes $\mu \in \mathbb{R}$ und $\sigma > 0$:

$$I = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1.$$

Beweis: Man betrachte:

$$I^2 = \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx\right) \cdot \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} dy\right).$$

Man mache die folgenden Substitutionen:

$$u = \frac{x-\mu}{\sigma} \quad \text{und} \quad v = \frac{y-\mu}{\sigma}.$$

Daraus folgt:

$$du = \frac{dx}{\sigma}, dv = \frac{dy}{\sigma} \text{ und daraus: } dx = \sigma du, dy = \sigma dv.$$

Die Integralgrenzen bleiben jeweils bei $-\infty$ und ∞. Somit erhalten wir:

$$\begin{aligned} I^2 &= \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}u^2} du\right) \cdot \left(\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}v^2} dv\right) \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}(u^2+v^2)} du\, dv. \end{aligned}$$

Nun machen wir eine Substitution mit Hilfe von Polarkoordinaten:

$$u = r\cos\phi, v = r\sin\phi \text{ mit } r > 0, \phi \in [0, 2\pi).$$

Die Jacobi-Determinante ist r. Somit erhalten wir:

$$\begin{aligned} I^2 &= \int_0^{2\pi} \int_0^{\infty} \frac{1}{2\pi} e^{-\frac{1}{2}r^2} \cdot r\, dr\, d\phi \\ &= \int_0^{2\pi} \frac{1}{2\pi} \left[-e^{-\frac{1}{2}r^2}\right]_0^{\infty} d\phi \\ &= \int_0^{2\pi} \frac{1}{2\pi} d\phi = 1. \end{aligned}$$

Da $I \geq 0$, muss also gelten: $I = 1$.

$\square$

Eigenschaften:

1) $f_X(x)$ ist spiegelsymmetrisch bezüglich $y = \mu$
2) f_X nimmt Maximum bei $x = \mu$ an. und Wendepunkte $x_{1,2} = \mu \pm \sigma$
3) f_X ist normiert (d.h ist eine Dichte $\int_{-\infty}^{+\infty} f_X(x)\,dx = 1$ → auf Folien bewiesen.
4) Während der Parameter μ die Lage des Max. festlegt, bestimmt σ Breite und Höhe der GlockenKurve. Je kleiner σ ist, desto höher liegt das max und umso steiler fällt die Dichtekurve nach beiden Seiten hin ab.
5) $\mathbb{P}[X \leq \mu] = \mathbb{P}[X \geq \mu] = \frac{1}{2}$

STANDARD NORMALVERTEILUNG.

- Sei $\boxed{X \sim \mathcal{N}(\mu, \sigma)}$ und $\boxed{Z = \frac{X-\mu}{\sigma}}$ (d.h Z ist eine FKt von X).

Dann • $F_Z(z) = \mathbb{P}[Z \leq z] = \mathbb{P}\left[\frac{X-\mu}{\sigma} \leq z\right] = \mathbb{P}[X \leq \sigma z + \mu] =$

$$= F_X(\sigma z + \mu)$$

- Dichte von Z.

$$f_Z(z) = \frac{\partial}{\partial z} F_Z(x) = \frac{\partial}{\partial z} F_X(\sigma z + \mu) = \sigma f_X(\sigma z + \mu) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2} \Rightarrow$$

$\Rightarrow$ Z ist Normalverteilt mit $\boxed{\mu = 0}$ und $\boxed{\sigma = 1}$

Def: Z heißt Standardnormalverteilt, falls die Dichtefkt

$$f_Z(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}$$

$$Z \sim \mathcal{N}(0,1)$$

Notation: Falls $Z \sim N(0,1)$, dann Verteilungsfkt von Z wird mit $\Phi(z) = \mathbb{P}[Z \leq z]$ bezeichnet.

Es gilt $\boxed{\Phi(-z) = 1 - \Phi(z)}$ Symmetrie!

Lemma [Kenngrößen von $Z \sim \mathcal{N}(0,1)$]

1) $\mathbb{E}[Z] = 0$
2) $Var(Z) = \sigma^2$

Beweis: 1) $\mathbb{E}[Z] = \int_{-\infty}^{+\infty} z \frac{1}{\sqrt{2\pi}} e^{-z^2/2} dz = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} \left(-e^{-\frac{z^2}{2}}\right)' dz =$

$$= \frac{1}{\sqrt{2\pi}} \left(-e^{-\frac{z^2}{2}}\right)\Big|_{-\infty}^{+\infty} = \frac{1}{\sqrt{2\pi}} (0-0) = 0$$

2) $Var(Z) = \mathbb{E}[Z^2] - (\underbrace{\mathbb{E}[Z]}_{0})^2$

$$\mathbb{E}[Z^2] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} z^2 e^{-\frac{z^2}{2}} dz = \frac{1}{\sqrt{2\pi}} \left(\underbrace{-z e^{-\frac{z^2}{2}} \Big|_{-\infty}^{+\infty}}_{0} + \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz \right) =$$

$$= \underbrace{\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{+\infty} e^{-\frac{z^2}{2}} dz}_{\text{Dichte einer Verteilung}} = 1$$

Lemma: Sei $X \sim \mathcal{N}(\mu, \sigma)$ und . Dann gilt

1) $\mathbb{E}[X] = \mu$
2) $Var(X) = \sigma^2$

Beweis: Da $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1) \Rightarrow \begin{array}{l} \mathbb{E}[Z] = 0 \\ Var(Z) = 1 \end{array}$

$$\Downarrow$$

$$\boxed{X = \sigma Z + \mu}$$

$$\mathbb{E}[X] = \sigma \underbrace{\mathbb{E}[Z]}_{0} + \mu = \mu$$

$$Var(X) = Var(\sigma Z + \mu) = \sigma^2 \underbrace{Var(Z)}_{1} = \sigma^2$$

Berechnung mit Wahrscheinlichkeiten.

0 über die Standardnormalverteilung

° Erinnerung: $X \sim \mathcal{N}(\mu,\sigma) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$

umgekehrt $X = \sigma Z + \mu \sim \mathcal{N}(\mu,\sigma)$

Verteilungsfkt: $\boxed{\Phi(z) = \mathbb{P}[Z \leq z]}$

Dann: $F_X(x) = \mathbb{P}[X \leq x] = \mathbb{P}[\sigma Z + \mu \leq x] = \mathbb{P}\left[Z \leq \frac{x-\mu}{\sigma}\right] =$

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

Bsp: $X \sim \mathcal{N}(1,2) \Rightarrow Z = \frac{X-1}{2} \sim \mathcal{N}(0,1)$

$$\mathbb{P}[-2 \leq X \leq 1] = \mathbb{P}[-2-1 \leq X-1 \leq 1-1] = \mathbb{P}\left[\frac{-3}{2} \leq \frac{X-1}{2} \leq 0\right] =$$

$$= \mathbb{P}\left[\frac{-3}{2} \leq Z \leq 0\right] = \underbrace{\Phi(0) - \Phi\left(-\frac{3}{2}\right)}_{\text{Tabelle der Standard N-Vtlg.}} = \Phi(0) - (1 - \Phi(1{,}5))$$

$$= 0{,}5 - 1 + 0{,}93 = 0{,}43$$

Seite 80

Satz: Sei $X \sim \mathcal{N}(\mu,\sigma)$ und $Z = \frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$. Dann.

1) $\mathbb{P}[-c \leq Z \leq c] = 2\Phi(c) - 1$

2) $\mathbb{P}[\mu - c \leq X \leq \mu + c] = 2\Phi\left(\frac{c}{\sigma}\right) - 1$

Beweis: 1) $\mathbb{P}[-c \leq Z \leq c] = \Phi(c) - \Phi(-c) = \Phi(c) - (1 - \Phi(c))$

$$= 2\Phi(c) - 1$$

2) $\mathbb{P}[\mu - c \leq X \leq \mu + c] = \mathbb{P}[\cancel{\mu} - c \leq \cancel{\mu} + \sigma Z \leq \cancel{\mu} + c] =$

$$= \mathbb{P}\left[\frac{-c}{\sigma} \leq Z \leq \frac{c}{\sigma}\right] = \Phi\left(\frac{c}{\sigma}\right) - \Phi\left(-\frac{c}{\sigma}\right) \overset{(1)}{=} 2\Phi\left(\frac{c}{\sigma}\right) - 1$$

Wie liest man die Tabelle?

$\Phi(0{,}83) = 0{,}7939$

$\Phi(1.44) = 0{,}9251$

$\Phi(2{,}68) = 0{,}9963$

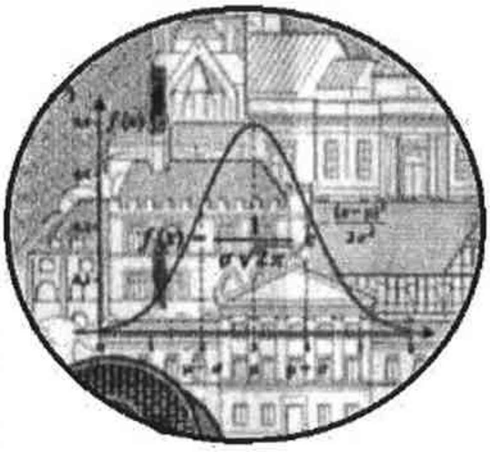
Tabelle der Normalverteilung

Tabelle des Integrals $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-t^2/2} dt$. Beispiel: $\Phi(1.23) = 0.8907$.

x	0	1	2	3	4	5	6	7	8	9
0.00	.5000	.5040	.5080	.5120	.5160	.5199	.5239	.5279	.5319	.5359
0.10	.5398	.5438	.5478	.5517	.5557	.5596	.5636	.5675	.5714	.5753
0.20	.5793	.5832	.5871	.5910	.5948	.5987	.6026	.6064	.6103	.6141
0.30	.6179	.6217	.6255	.6293	.6331	.6368	.6406	.6443	.6480	.6517
0.40	.6554	.6591	.6628	.6664	.6700	.6736	.6772	.6808	.6844	.6879
0.50	.6915	.6950	.6985	.7019	.7054	.7088	.7123	.7157	.7190	.7224
0.60	.7257	.7291	.7324	.7357	.7389	.7422	.7454	.7486	.7517	.7549
0.70	.7580	.7611	.7642	.7673	.7704	.7734	.7764	.7794	.7823	.7852
0.80	.7881	.7910	.7939	.7967	.7995	.8023	.8051	.8079	.8106	.8133
0.90	.8159	.8186	.8212	.8238	.8264	.8289	.8315	.8340	.8365	.8389
1.00	.8413	.8438	.8461	.8485	.8508	.8531	.8554	.8577	.8599	.8621
1.10	.8643	.8665	.8686	.8708	.8729	.8749	.8770	.8790	.8810	.8830
1.20	.8849	.8869	.8888	.8907	.8925	.8944	.8962	.8980	.8997	.9015
1.30	.9032	.9049	.9066	.9082	.9099	.9115	.9131	.9147	.9162	.9177
1.40	.9192	.9207	.9222	.9236	.9251	.9265	.9279	.9292	.9306	.9319
1.50	.9332	.9345	.9357	.9370	.9382	.9394	.9406	.9418	.9429	.9441
1.60	.9452	.9463	.9474	.9485	.9495	.9505	.9515	.9525	.9535	.9545
1.70	.9554	.9564	.9573	.9582	.9591	.9599	.9608	.9616	.9625	.9633
1.80	.9641	.9649	.9656	.9664	.9671	.9678	.9686	.9693	.9699	.9706
1.90	.9713	.9719	.9726	.9732	.9738	.9744	.9750	.9756	.9762	.9767
2.00	.9773	.9778	.9783	.9788	.9793	.9798	.9803	.9808	.9812	.9817
2.10	.9821	.9826	.9830	.9834	.9838	.9842	.9846	.9850	.9854	.9857
2.20	.9861	.9865	.9868	.9871	.9875	.9878	.9881	.9884	.9887	.9890
2.30	.9893	.9896	.9898	.9901	.9904	.9906	.9909	.9911	.9913	.9916
2.40	.9918	.9920	.9922	.9925	.9927	.9929	.9931	.9932	.9934	.9936
2.50	.9938	.9940	.9941	.9943	.9945	.9946	.9948	.9949	.9951	.9952
2.60	.9953	.9955	.9956	.9957	.9959	.9960	.9961	.9962	.9963	.9964
2.70	.9965	.9966	.9967	.9968	.9969	.9970	.9971	.9972	.9973	.9974
2.80	.9974	.9975	.9976	.9977	.9977	.9978	.9979	.9980	.9980	.9981
2.90	.9981	.9982	.9983	.9983	.9984	.9984	.9985	.9985	.9986	.9986

Folie zur Vorlesung
"Wahrscheinlichkeitsrechnung und Stoch. Prozesse"

Aufgaben zur Normalverteilung:



In einer Schokoladenfabrik werden Schokoladentafeln produziert. Das Gewicht X einer zufällig ausgewählten Tafel sei normal-verteilt mit Parameter $\mu = 100$ Gramm und $\sigma = 1,5$.

1. Wie groß ist die Wahrscheinlichkeit, daß eine zufällig ausgewählte Tafel ein Gewicht zwischen 98 und 101 Gramm besitzt?
2. Was ist mit mindestens 99% Wahrscheinlichkeit das Mindestgewicht einer zufällig ausgewählten Tafel Schokolade?
3. Bestimmen Sie ein um μ symmetrisches Intervall, in welchem zu 95% das Gewicht einer zufällig ausgewählten Schokoladentafel liegt.

1) $X \sim \mathcal{N}(100, 1{,}5) \Rightarrow Z = \frac{X-100}{1{,}5} \sim \mathcal{N}(0,1)$

$\mathbb{P}[98 \le X \le 101] = \mathbb{P}\left[\underbrace{\frac{98-100}{1{,}5}}_{-1{,}33} \le Z \le \underbrace{\frac{101-100}{1{,}5}}_{0{,}67}\right] = \Phi(0{,}67) - \Phi(-1{,}33) =$

$= \Phi(0{,}67) - (1 - \Phi(1{,}33)) =$

$= 0{,}7486 - 1 + 0{,}9082 = 0{,}6568$

2) $\mathbb{P}[X \ge c] = 99\% \Rightarrow \mathbb{P}\left[\underbrace{\frac{X-100}{1{,}5}}_{Z} \ge \frac{c-100}{1{,}5}\right] = 99\% \Rightarrow 1 - \mathbb{P}\left[Z \le \frac{c-100}{1{,}5}\right]$

$= 99\% \Rightarrow \boxed{\Phi\left(\frac{c-100}{1{,}5}\right) = \frac{1}{100}}$

$\Phi(2{,}33) = 0{,}99$

$\Phi(-2{,}33) = 1 - \Phi(2{,}33) = 0{,}01$

$\Rightarrow \frac{c-100}{1{,}5} = -2{,}33 \Rightarrow \boxed{c = 96{,}505}$

3) $\mathbb{P}[100+c \le X \le 100-c] = 0{,}95 \Rightarrow \Rightarrow \mathbb{P}\left[\frac{c}{1{,}5} \le \frac{X-100}{1{,}5} \le \frac{c}{1{,}5}\right] = 0{,}95 \Rightarrow$

$2\Phi\left(\frac{c}{1{,}5}\right) - 1 = 0{,}95 \Rightarrow \Phi\left(\frac{c}{1{,}5}\right) = 0{,}975 \Rightarrow \frac{c}{1{,}5} = 1{,}96 \Rightarrow \boxed{c = 2{,}94}$

7.3 Approximationen

Satz [Grenzwertsatz De Moivre-Laplace] → (auf Folien)

Sei $X \sim \mathcal{B}(n,p)$

Für n groß genug, i.e wenn $n \min\{p, 1-p\} > 5$

1) Für $k \in \{0,1,\dots,n\}$

$$\mathbb{P}[X=k] \approx \underbrace{\frac{1}{\sqrt{2\pi\, np(1-p)}} \cdot e^{-\frac{(k-np)^2}{2np(1-p)}}}_{\text{Dichte von } Y \sim \mathcal{N}(np, \sqrt{np(1-p)})}$$

2) Für $k, l \in \{0,1,\dots,n\}$, $l \le k$

$$\mathbb{P}[l \le X \le k] \approx \Phi\left(\frac{k+\frac{1}{2}-np}{\sqrt{np(1-p)}}\right) - \Phi\left(\frac{l-\frac{1}{2}-np}{\sqrt{np(1-p)}}\right)$$

Bsp $X \sim \mathcal{B}\left(2000, \frac{18}{37}\right)$

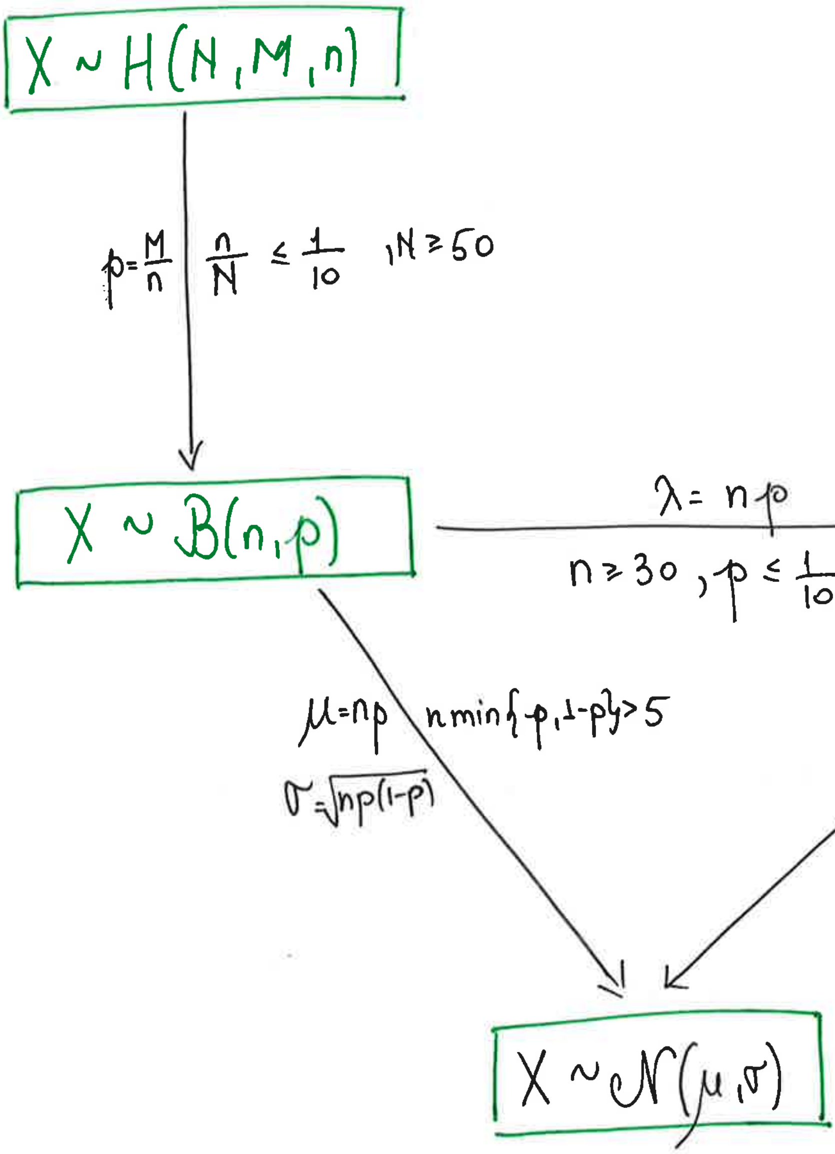
$$\mathbb{P}[980 \le X \le 1020] \approx \Phi\left(\frac{1020+\frac{1}{2}-972{,}97}{\sqrt{499{,}6}}\right) - \Phi\left(\frac{980-\frac{1}{2}-972{,}97}{\sqrt{499{,}6}}\right) =$$

$$np = 2000 \cdot \frac{18}{37} = 972{,}97$$

$$np(1-p) = 2000 \cdot \frac{18}{37} \cdot \frac{19}{37} = 499{,}6$$

$$= \Phi(2{,}13) - \Phi(0{,}29) = 0{,}3693$$

Bemerkung!: Die W-Funktion der Binomialverteilung $\mathcal{B}(n,p)$ wird durch die Dichte der Normalverteilung $\mathcal{N}(mp, \sqrt{np(1-p)})$ approximiert.


$X \sim H(N, M, n)$
$p = \frac{M}{n}$ $\frac{n}{N} \leq \frac{1}{10}$, $N \geq 50$
$X \sim B(n, p)$
$\lambda = np$
$n \geq 30$, $p \leq \frac{1}{10}$
$X \sim P(\lambda)$
$\mu = np$ $n \min\{p, 1-p\} > 5$
$\sigma = \sqrt{np(1-p)}$
$\lambda > 9$ $\mu = \lambda$
$\sigma = \sqrt{\lambda}$
$X \sim N(\mu, \sigma)$

(84)

Folie zur Vorlesung
"Wahrscheinlichkeitsrechnung und Stoch. Prozesse"
01.12.2016

Schätzung von Parametern der Normalverteilung:

Der zufällige Meßfehler X eines Meßgerätes werde als normal-verteilt angenommen mit unbekanntem Mittelwert μ und unbekannter Varianz σ^2. Es werden n unabhängige Probemessungen durchgeführt und jeweils die Abweichung vom tatsächlichen Wert (z.B. mittels Gegenprüfung durch ein geeichtes Gerät) betrachtet.

Beobachtete Werte: $x_1, x_2, \ldots, x_{n-1}, x_n$.

Der Mittelwert μ kann geschätzt werden durch

$$\bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i.$$

Die Varianz σ^2 kann geschätzt werden durch

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (x_i - \bar{x})^2.$$

Beispiel: 10 Messungen ergaben folgende Messfehler(Abweichungen vom tatsächlichen Wert):

$$-0.7,\ -0.5,\ -0.3,\ -0.2,\ 0,\ 0.1,\ 0.1,\ 0.2,\ 0.3,\ 0.8.$$

Wie groß ist die Wahrscheinlichkeit, daß eine Messung um maximal 0.5 vom tatsächlichen Wert abweicht?

Schätzungen für μ und σ^2:

$$\begin{aligned}
\mu &= \frac{1}{10}\big(-0.7 + (-0.5) + \ldots + 0.3 + 0.8\big) = -0.02, \\
\sigma^2 &= \frac{1}{9}\big((-0.7 - (-0.02))^2 + (-0.5 - (-0.02))^2 + \ldots + (0.8 - (-0.02))^2\big) = 0.203.
\end{aligned}$$

D.h. σ wird geschätzt durch $\sqrt{0.203}$. Der zufällige Messfehler X wird also durch eine Normalverteilung mit Parameter $\mu = -0.02$ und $\sigma = \sqrt{0.203}$ modelliert. Wir erhalten:

$$\begin{aligned}
\mathbb{P}[-0.5 \leq X \leq 0.5] &= \Phi\Big(\frac{0.5 - (-0.02)}{\sqrt{0.203}}\Big) - \Phi\Big(\frac{-0.5 - (-0.02)}{\sqrt{0.203}}\Big) \\
&= \Phi(1.15) - \Phi(-1.07) \\
&= \Phi(1.15) - 1 + \Phi(1.07) \\
&= 0.8749 - 1 + 0.8577 = 0.7326
\end{aligned}$$