Quantile von Zufallsvariablen I

Definition 9.11 (p-Quantil)

Seien X eine eindimensionale Zufallsvariable, $p \in (0,1)$. Jeder Wert $x_p \in \mathbb{R}$ mit

$$P\{X \le x_p\} \ge p \qquad \text{und} \qquad P\{X \ge x_p\} \ge 1-p$$

*heißt p-**Quantil** (auch p-**Perzentil**) von X. Man nennt Werte x_p mit dieser Eigenschaft spezieller*

- ***Median** von X für $p = 0.5$,*
- ***unteres Quartil** von X für $p = 0.25$ sowie*
- ***oberes Quartil** von X für $p = 0.75$.*

Ist F_X die Verteilungsfunktion von X, so ist x_p also genau dann p-Quantil von X, wenn

$$F_X(x_p - 0) \le p \le F_X(x_p)$$

gilt, für stetige Zufallsvariablen X also genau dann, wenn $F_X(x_p) = p$ gilt.

Quantile von Zufallsvariablen II

- p-Quantile sind nach Definition 9.11 eindeutig bestimmt, wenn die Verteilungsfunktion F_X der Zufallsvariablen X (dort, wo sie Werte in $(0,1)$ annimmt) invertierbar ist, also insbesondere stetig und streng monoton wachsend.
  Bezeichnet F_X^{-1} die Umkehrfunktion von F_X, so gilt dann
  $$x_p = F_X^{-1}(p) \qquad \text{für alle } p \in (0,1)\ .$$
  F_X^{-1} wird in diesem Fall auch **Quantilsfunktion** genannt.
- Der Abstand $x_{0.75} - x_{0.25}$ zwischen unterem und oberen Quartil wird (wie auch bei empirischen Quartilen) auch **Interquartilsabstand (IQA)** genannt.

Quantile von Zufallsvariablen III

- Auch ohne die Invertierbarkeit von F_X kann Eindeutigkeit von Quantilen zum Beispiel durch die Festsetzung
  $$x_p := \min\{x \mid P\{X \le x\} \ge p\} = \min\{x \mid F_X(x) \ge p\}$$
  erreicht werden.
  Man nennt die Abbildung
  $$(0,1) \to \mathbb{R}; p \mapsto x_p = \min\{x \mid F_X(x) \ge p\}$$
  häufig auch *verallgemeinerte Inverse* von F_X und verwendet hierfür dann ebenfalls das Symbol F_X^{-1} sowie die Bezeichnung Quantilsfunktion.
- Diese Eindeutigkeitsfestlegung **unterscheidet** sich von der vergleichbaren Konvention aus der deskriptiven Statistik für empirische Quantile!

Spezielle diskrete Verteilungen

- Im Folgenden: Vorstellung spezieller (parametrischer) **Verteilungsfamilien**, die häufig Verwendung finden.
- Häufige Verwendung ist dadurch begründet, dass diese Verteilungen in vielen verschiedenen Anwendungen anzutreffen sind bzw. die Zufallsabhängigkeit interessanter Größen geeignet modellieren.
- Parametrische Verteilungsfamilien sind Mengen von (ähnlichen) Verteilungen Q_θ, deren Elemente sich nur durch die Ausprägung eines oder mehrerer **Verteilungsparameter** unterscheiden, d.h. die spezielle Verteilung hängt von einem Parameter oder einem Parametervektor θ ab, und zu jedem Parameter(vektor) gehört jeweils eine eigene Verteilung Q_θ.
- Die Menge aller möglichen Parameter(vektoren) θ, auch **Parameterraum** genannt, wird meist mit Θ bezeichnet. Die Verteilungsfamilie ist damit die Menge $\{Q_\theta \mid \theta \in \Theta\}$.
- Besitzt eine Zufallsvariable X die Verteilung Q_θ, so schreibt man auch kurz: $X \sim Q_\theta$.
- *Zunächst:* Vorstellung spezieller *diskreter* Verteilungen.

Bernoulli-/Alternativverteilung

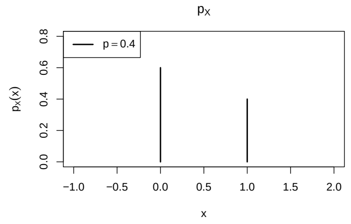
- Verwendung:
  - Modellierung eines Zufallsexperiments $(\Omega, \mathcal{F}, P)$, in dem nur das Eintreten bzw. Nichteintreten eines einzigen Ereignisses A von Interesse ist.
  - Eintreten des Ereignisses A wird oft als „Erfolg" interpretiert, Nichteintreten (bzw. Eintreten von $\overline{A}$) als „Misserfolg".
  - Zufallsvariable soll im Erfolgsfall Wert 1 annehmen, im Misserfolgsfall Wert 0, es sei also
    $$X(\omega) := \begin{cases} 1 & \text{falls } \omega \in A \\ 0 & \text{falls } \omega \in \overline{A} \end{cases}$$
  - Beispiel: Werfen eines fairen Würfels, Ereignis A: „6 gewürfelt" mit $P(A) = \frac{1}{6}$.
- Verteilung von X hängt damit *nur* von „Erfolgswahrscheinlichkeit" $p := P(A)$ ab; p ist also einziger Parameter der Verteilungsfamilie.
- Um triviale Fälle auszuschließen, betrachtet man nur Ereignisse mit $p \in (0,1)$
- Der Träger der Verteilung ist dann $T(X) = \{0,1\}$, die Punktwahrscheinlichkeiten sind $p_X(0) = 1 - p$ und $p_X(1) = p$.
- Symbolschreibweise für Bernoulli-Verteilung mit Parameter p: $B(1,p)$
- Ist X also Bernoulli-verteilt mit Parameter p, so schreibt man $X \sim B(1,p)$.

Bernoulli-/Alternativverteilung
$B(1,p)$

Parameter:
$p \in (0,1)$

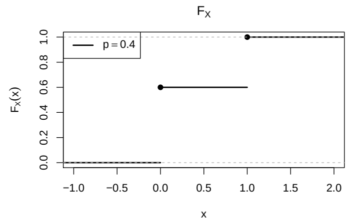
Träger: $T(X) = \{0,1\}$
Wahrscheinlichkeitsfunktion:

$$p_X(x) = \begin{cases} 1-p & \text{für } x = 0 \\ p & \text{für } x = 1 \\ 0 & \text{sonst} \end{cases}$$



Verteilungsfunktion:

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1-p & \text{für } 0 \leq x < 1 \\ 1 & \text{für } x \geq 1 \end{cases}$$



Momente: $E(X) = p$ $\quad$ $\mathrm{Var}(X) = p \cdot (1-p)$

$\gamma(X) = \frac{1-2p}{\sqrt{p(1-p)}}$ $\quad$ $\kappa(X) = \frac{1-3p(1-p)}{p(1-p)}$

Binomialverteilung

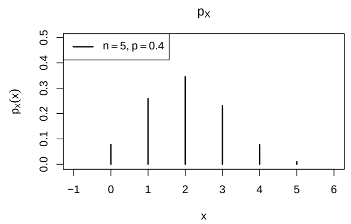
- Verallgemeinerung der Bernoulli-Verteilung
- Verwendung:
  - Modellierung der **unabhängigen, wiederholten** Durchführung eines Zufallsexperiments, in dem nur die **Häufigkeit** des Eintretens bzw. Nichteintretens eines Ereignisses A interessiert („Bernoulli-Experiment").
  - Eintreten des Ereignisses A wird auch hier oft als „Erfolg" interpretiert, Nichteintreten (bzw. Eintreten von $\overline{A}$) als „Misserfolg".
  - Zufallsvariable X soll die **Anzahl der Erfolge** bei einer vorgegebenen Anzahl von n Wiederholungen des Experiments zählen.
  - Nimmt X_i für $i \in \{1, \ldots, n\}$ im Erfolgsfall (für Durchführung i) den Wert 1 an, im Misserfolgsfall den Wert 0, dann gilt also $X = \sum_{i=1}^{n} X_i$.
  - Beispiel: 5-faches Werfen eines fairen Würfels, Anzahl der Zahlen kleiner 3. $\rightsquigarrow n = 5, p = 1/3$.
- Verteilung von X hängt damit *nur* von „Erfolgswahrscheinlichkeit" $p := P(A)$ sowie der Anzahl der Durchführungen n des Experiments ab.
- Um triviale Fälle auszuschließen, betrachtet man nur die Fälle $n \in \mathbb{N}$ und $p \in (0,1)$. Träger der Verteilung ist dann $T(X) = \{0, 1, \ldots, n\}$.
- Symbolschreibweise für Binomialverteilung mit Parameter n und p: $B(n,p)$
- Übereinstimmung mit Bernoulli-Verteilung (mit Parameter p) für $n = 1$.

Binomialverteilung
$B(n,p)$

Parameter:
$n \in \mathbb{N}, p \in (0,1)$

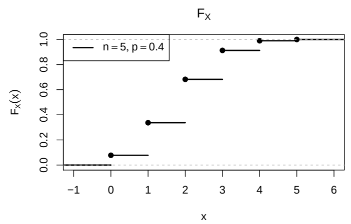
Träger: $T(X) = \{0, 1, \ldots, n\}$
Wahrscheinlichkeitsfunktion: $p_X(x)$

$$= \begin{cases} \binom{n}{x} p^x (1-p)^{n-x} & \text{für } x \in T(X) \\ 0 & \text{sonst} \end{cases}$$



Verteilungsfunktion:

$$F_X(x) = \sum_{\substack{x_i \in T(X) \\ x_i \leq x}} p_X(x_i)$$



Momente: $E(X) = n \cdot p$ $\quad$ $\mathrm{Var}(X) = n \cdot p \cdot (1-p)$

$\gamma(X) = \frac{1-2p}{\sqrt{np(1-p)}}$ $\quad$ $\kappa(X) = \frac{1+(3n-6)p(1-p)}{np(1-p)}$

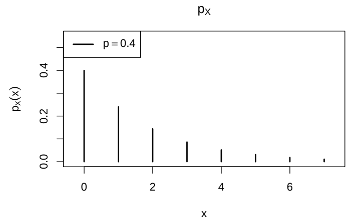
Geometrische Verteilung

- Verwendung:
  - Modellierung der **unabhängigen, wiederholten** Durchführung eines Bernoulli-Experiments (nur das Eintreten bzw. Nichteintreten eines einzigen Ereignisses A ist von Interesse), bis das Ereignis A **zum ersten Mal** eintritt.
  - Zufallsvariable X zählt **Anzahl der Misserfolge**, ausschließlich des (letzten) „erfolgreichen" Versuchs, bei dem Ereignis A zum ersten Mal eintritt.
  - X kann also nur Werte $x \in \mathbb{N}_0$ annehmen, man erhält die Realisation x von X, wenn nach genau x Misserfolgen (Nicht-Eintreten von A) in der $(x+1)$-ten Durchführung ein Erfolg (Eintreten von A) zu verzeichnen ist.
  - Ist $p := P(A)$ die „Erfolgswahrscheinlichkeit" des Bernoulli-Experiments, so gilt offensichtlich $P\{X = x\} = (1-p)^x \cdot p$ für alle $x \in \mathbb{N}_0$.
  - Beispiel (vgl. Folie 168): Anzahl des Auftretens von „Zahl" beim Werfen einer Münze („Wappen" oder „Zahl"), bis zum ersten Mal „Wappen" erscheint $\rightsquigarrow p = 1/2$ (bei fairer Münze).
- Verteilung von X hängt damit *nur* von Erfolgswahrscheinlichkeit p ab.
- Um triviale Fälle auszuschließen, betrachtet man nur den Fall $p \in (0,1)$. Träger der Verteilung ist dann $T(X) = \mathbb{N}_0 = \{0, 1, \ldots\}$.
- Symbolschreibweise für geometrische Verteilung mit Parameter p: Geom(p)

Geometrische Verteilung	Parameter:
Geom(p)	$p \in (0,1)$

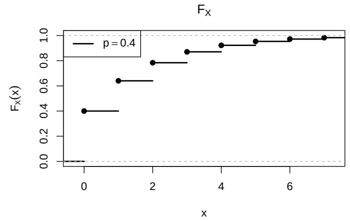
Träger: $T(X) = \mathbb{N}_0 = \{0, 1, \ldots\}$
Wahrscheinlichkeitsfunktion:

$$p_X(x) = \begin{cases} (1-p)^x \cdot p & \text{für } x \in T(X) \\ 0 & \text{sonst} \end{cases}$$



Verteilungsfunktion:

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1-(1-p)^{\lfloor x \rfloor + 1} & \text{für } x \geq 0 \end{cases}$$



Momente: $E(X) = \frac{1-p}{p}$ $\quad$ $\mathrm{Var}(X) = \frac{1-p}{p^2}$
$\gamma(X) = \frac{2-p}{\sqrt{1-p}}$ $\quad$ $\kappa(X) = \frac{p^2-9p+9}{1-p}$

Poisson-Verteilung

- „Grenzverteilung" der Binomialverteilung
- Verwendung:
  - Approximation einer $B(n,p)$-Verteilung, wenn n (sehr) groß und p (sehr) klein ist.
  - „Faustregeln" zur Anwendung der Approximation:

$$n \geq 50, \qquad p \leq 0.1, \qquad n \cdot p \leq 10$$

  - Poisson-Verteilung hat einzigen Parameter $\lambda > 0$, der zur Approximation einer $B(n,p)$-Verteilung auf $\lambda = n \cdot p$ gesetzt wird.
- Träger von Poisson-verteilten Zufallsvariablen X: $T(X) = \mathbb{N}_0 = \{0, 1, \ldots\}$
- Wahrscheinlichkeitsfunktion für $x \in T(X)$: $p_X(x) = \frac{\lambda^x}{x!} e^{-\lambda}$, wobei $e = \exp(1)$ die Eulersche Zahl ist, also $e \approx 2.71828$.
- Gültigkeit der Approximation beruht auf Konvergenz der Punktwahrscheinlichkeiten. Es gilt nämlich für alle $x \in \mathbb{N}_0$:

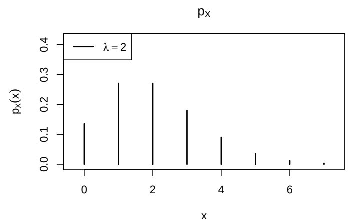
$$\lim_{\substack{n \to \infty \\ p \to 0 \\ n \cdot p \to \lambda}} \binom{n}{x} p^x (1-p)^{n-x} = \frac{\lambda^x}{x!} e^{-\lambda}$$

- Symbolschreibweise für Poisson-Verteilung mit Parameter λ: Pois(λ)

Poisson-Verteilung	Parameter:
Pois(λ)	$\lambda > 0$

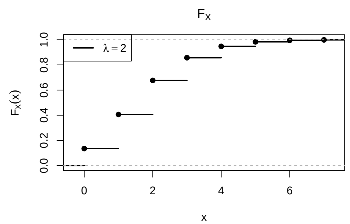
Träger: $T(X) = \mathbb{N}_0 = \{0, 1, \ldots\}$
Wahrscheinlichkeitsfunktion:

$$p_X(x) = \begin{cases} \frac{\lambda^x}{x!} e^{-\lambda} & \text{für } x \in T(X) \\ 0 & \text{sonst} \end{cases}$$



Verteilungsfunktion:

$$F_X(x) = \sum_{\substack{x_i \in T(X) \\ x_i \leq x}} p_X(x_i)$$



Momente: $E(X) = \lambda$ $\quad$ $\mathrm{Var}(X) = \lambda$
$\gamma(X) = \frac{1}{\sqrt{\lambda}}$ $\quad$ $\kappa(X) = 3 + \frac{1}{\lambda}$

Spezielle stetige Verteilungen

- Nun: Vorstellung spezieller parametrischer Verteilungsfamilien von *stetigen* Verteilungen
- In Verallgemeinerung des **Trägers** diskreter Verteilungen:
  Träger $T(X)$ einer stetigen Verteilung als **„Bereich positiver Dichte“**.
- Wegen Möglichkeit, Dichtefunktionen abzuändern, etwas genauer:

$$\begin{aligned} T(X) := \{x \in \mathbb{R} \,|\, &\text{es gibt eine Dichtefunktion } f_X \text{ von } X \text{ und ein } \epsilon > 0 \\ &\text{mit } (f_X(t) > 0 \text{ für alle } t \in [x-\epsilon, x]) \\ &\text{oder } (f_X(t) > 0 \text{ für alle } t \in [x, x+\epsilon])\} \end{aligned}$$

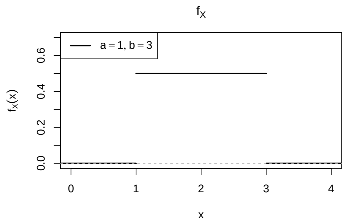
Stetige Gleichverteilung

- Einfachste stetige Verteilungsfamilie:
  Stetige Gleichverteilung auf Intervall $[a, b]$
- Modellierung einer stetigen Verteilung, in der alle Realisationen in einem Intervall $[a, b]$ als „gleichwahrscheinlich“ angenommen werden.
- Verteilung hängt von den beiden Parametern $a, b \in \mathbb{R}$ mit $a < b$ ab.
- Dichtefunktion f_X einer gleichverteilten Zufallsvariablen X kann auf Intervall $[a, b]$ konstant zu $\frac{1}{b-a}$ gewählt werden.
- Träger der Verteilung: $T(X) = [a, b]$
- Symbolschreibweise für stetige Gleichverteilung auf $[a, b]$: $X \sim \text{Unif}(a, b)$

Stetige Gleichverteilung Unif(a, b)	Parameter: $a, b \in \mathbb{R}$ mit $a < b$

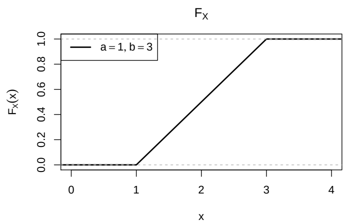
Träger: $T(X) = [a, b]$
Dichtefunktion: $f_X : \mathbb{R} \to \mathbb{R}$;

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{für } a \le x \le b \\ 0 & \text{sonst} \end{cases}$$



Verteilungsfunktion: $F_X : \mathbb{R} \to \mathbb{R}$;

$$F_X(x) = \begin{cases} 0 & \text{für } x < a \\ \frac{x-a}{b-a} & \text{für } a \le x \le b \\ 1 & \text{für } x > b \end{cases}$$



Momente: $E(X) = \frac{a+b}{2}$ $\quad \text{Var}(X) = \frac{(b-a)^2}{12}$
$\gamma(X) = 0$ $\quad \kappa(X) = \frac{9}{5}$

Normalverteilung

- Verteilung entsteht als Grenzverteilung bei Durchschnittsbildung vieler (unabhängiger) Zufallsvariablen (später mehr!) $\rightsquigarrow$ Einsatz für Näherungen
- Familie der Normalverteilungen hat Lageparameter $\mu \in \mathbb{R}$, der mit Erwartungswert übereinstimmt, und Streuungsparameter $\sigma^2 > 0$, der mit Varianz übereinstimmt, Standardabweichung ist dann $\sigma := +\sqrt{\sigma^2}$.
- Verteilungsfunktion von Normalverteilungen schwierig zu handhaben, Berechnung muss i.d.R. mit Software/Tabellen erfolgen.
- Wichtige Eigenschaft der Normalverteilungsfamilie:
  Ist X normalverteilt mit Parameter $\mu = 0$ und $\sigma^2 = 1$, dann ist $aX + b$ für $a, b \in \mathbb{R}$ normalverteilt mit Parameter $\mu = b$ und $\sigma^2 = a^2$.

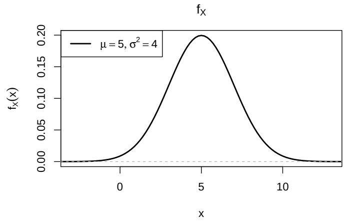
$\rightsquigarrow$ Zurückführung allgemeiner Normalverteilungen auf den Fall der **Standardnormalverteilung (Gauß-Verteilung)** mit Parameter $\mu = 0$ und $\sigma^2 = 1$, Tabellen/Algorithmen für Standardnormalverteilung damit einsetzbar.

- Dichtefunktion der Standardnormalverteilung: φ, Verteilungsfunktion: Φ.
- Träger aller Normalverteilungen ist $T(X) = \mathbb{R}$.
- Symbolschreibweise für Normalverteilung mit Parameter μ, σ^2: $X \sim \text{N}(\mu, \sigma^2)$

Normalverteilung	Parameter:
$N(\mu, \sigma^2)$	$\mu \in \mathbb{R}, \sigma^2 > 0$

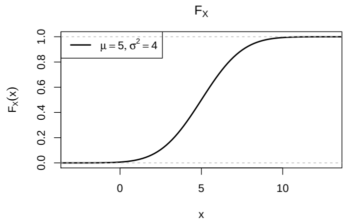
Träger: $T(X) = \mathbb{R}$
Dichtefunktion: $f_X : \mathbb{R} \to \mathbb{R}$;

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}} = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right)$$



Verteilungsfunktion:

$$F_X : \mathbb{R} \to \mathbb{R}; F_X(x) = \Phi\left(\frac{x-\mu}{\sigma}\right)$$



Momente: $E(X) = \mu$ $\quad$ $\mathrm{Var}(X) = \sigma^2$
$\gamma(X) = 0$ $\quad$ $\kappa(X) = 3$

Exponentialverteilung

- Beliebte Verteilungsfamilie zur Modellierung von **Wartezeiten**.
- Verteilung entsteht als Grenzverteilung der geometrischen Verteilung (Anzahl Fehlversuche vor erstem Erfolg bei wiederholter, unabhängiger Ausführung eines Bernoulli-Experiments) bei Erfolgswahrscheinlichkeit $p \to 0$.
- Da die Anzahl X der benötigten Versuche für $p \to 0$ offensichtlich immer größere Werte annehmen wird, wird statt der *Anzahl* der benötigten Versuche die *Zeit* zur Durchführung der benötigten Versuche modelliert, und mit $p \to 0$ zugleich die *pro Zeiteinheit* durchgeführten Versuche n des Bernoulli-Experiments so erhöht, dass $p \cdot n =: \lambda$ konstant bleibt.
- Einziger Parameter der resultierenden Exponentialverteilung ist damit die als „erwartete Anzahl von Erfolgen pro Zeiteinheit" interpretierbare Größe $\lambda > 0$.
- Ist X exponentialverteilt mit Parameter λ, so erhält man $F_X(x)$ aus der Verteilungsfunktion der geometrischen Verteilung für $x \geq 0$ gemäß

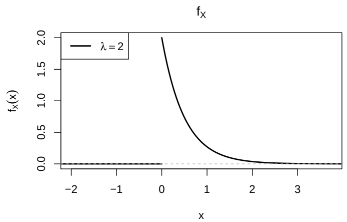
$$F_X(x) = \lim_{n\to\infty} 1 - \left(1 - \frac{\lambda}{n}\right)^{n\cdot x} = \lim_{n\to\infty} 1 - \left(1 + \frac{-\lambda\cdot x}{n\cdot x}\right)^{n\cdot x} = 1 - e^{-\lambda x}\,.$$

- Träger der Exponentialverteilungsfamilie ist $\mathbb{R}_+ := \{x \in \mathbb{R} \,|\, x \geq 0\}$.
- Symbolschreibweise für Exponentialverteilung mit Parameter λ: $X \sim \mathrm{Exp}(\lambda)$

Exponentialverteilung	Parameter:
$\mathrm{Exp}(\lambda)$	$\lambda > 0$

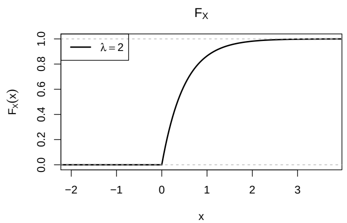
Träger: $T(X) = \mathbb{R}_+$
Dichtefunktion: $f_X : \mathbb{R} \to \mathbb{R}$;

$$f_X(x) = \begin{cases} \lambda \cdot e^{-\lambda x} & \text{für } x \geq 0 \\ 0 & \text{sonst} \end{cases}$$



Verteilungsfunktion: $F_X : \mathbb{R} \to \mathbb{R}$;

$$F_X(x) = \begin{cases} 0 & \text{für } x < 0 \\ 1 - e^{-\lambda x} & \text{für } x \geq 0 \end{cases}$$



Momente: $E(X) = \frac{1}{\lambda}$ $\quad$ $\mathrm{Var}(X) = \frac{1}{\lambda^2}$
$\gamma(X) = 2$ $\quad$ $\kappa(X) = 9$

Verwendung spezieller Verteilungen

- Übliche Vorgehensweise zur Berechnung von (Intervall-)Wahrscheinlichkeiten für Zufallsvariablen X: Verwendung der Verteilungsfunktion F_X
- Problem bei einigen der vorgestellten Verteilungen:
  Verteilungsfunktion F_X schlecht handhabbar bzw. nicht leicht auszuwerten!
- Traditionelle Lösung des Problems: *Vertafelung* bzw. *Tabellierung* der Verteilungsfunktionswerte, Ablesen der Werte dann aus Tabellenwerken.
- Lösung nicht mehr zeitgemäß: (kostenlose) PC-Software für alle benötigten Verteilungsfunktionen verfügbar, zum Beispiel Statistik-Software **R** (`http://www.r-project.org`)
- **Aber:** In Klausur keine PCs verfügbar, daher dort Rückgriff auf Tabellen.
- Problematische Verteilungsfunktionen (bisher) sind die der Standardnormalverteilung, Binomialverteilung sowie Poisson-Verteilung.
- Tabellen oder Tabellenausschnitte zu diesen Verteilungen werden in Klausur (sofern benötigt) zur Verfügung gestellt!
- Auch das Bestimmen von Quantilen ist für diese Verteilungen nicht ohne Hilfsmittel möglich und muss mit Hilfe weiterer Tabellen oder auf Grundlage der tabellierten Verteilungsfunktionswerte erfolgen.

Ausschnitt aus Tabelle für $\Phi(x)$

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	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	0.5753
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	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	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	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	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	0.9864	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	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936

Ausschnitt aus Tabelle für $F_{B(n,p)}(x)$

n	x	$p = 0.05$	$p = 0.10$	$p = 0.15$	$p = 0.20$	$p = 0.25$	$p = 0.30$	$p = 0.35$	$p = 0.40$
1	0	0.9500	0.9000	0.8500	0.8000	0.7500	0.7000	0.6500	0.6000
1	1	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
2	0	0.9025	0.8100	0.7225	0.6400	0.5625	0.4900	0.4225	0.3600
2	1	0.9975	0.9900	0.9775	0.9600	0.9375	0.9100	0.8775	0.8400
2	2	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
3	0	0.8574	0.7290	0.6141	0.5120	0.4219	0.3430	0.2746	0.2160
3	1	0.9928	0.9720	0.9392	0.8960	0.8438	0.7840	0.7182	0.6480
3	2	0.9999	0.9990	0.9966	0.9920	0.9844	0.9730	0.9571	0.9360
3	3	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
4	0	0.8145	0.6561	0.5220	0.4096	0.3164	0.2401	0.1785	0.1296
4	1	0.9860	0.9477	0.8905	0.8192	0.7383	0.6517	0.5630	0.4752
4	2	0.9995	0.9963	0.9880	0.9728	0.9492	0.9163	0.8735	0.8208
4	3	1.0000	0.9999	0.9995	0.9984	0.9961	0.9919	0.9850	0.9744
4	4	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
5	0	0.7738	0.5905	0.4437	0.3277	0.2373	0.1681	0.1160	0.0778
5	1	0.9774	0.9185	0.8352	0.7373	0.6328	0.5282	0.4284	0.3370
5	2	0.9988	0.9914	0.9734	0.9421	0.8965	0.8369	0.7648	0.6826
5	3	1.0000	0.9995	0.9978	0.9933	0.9844	0.9692	0.9460	0.9130
5	4	1.0000	1.0000	0.9999	0.9997	0.9990	0.9976	0.9947	0.9898
5	5	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
6	0	0.7351	0.5314	0.3771	0.2621	0.1780	0.1176	0.0754	0.0467
6	1	0.9672	0.8857	0.7765	0.6554	0.5339	0.4202	0.3191	0.2333
6	2	0.9978	0.9842	0.9527	0.9011	0.8306	0.7443	0.6471	0.5443
6	3	0.9999	0.9987	0.9941	0.9830	0.9624	0.9295	0.8826	0.8208
6	4	1.0000	0.9999	0.9996	0.9984	0.9954	0.9891	0.9777	0.9590
6	5	1.0000	1.0000	1.0000	0.9999	0.9998	0.9993	0.9982	0.9959
6	6	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

R-Befehle für spezielle Verteilungen

- Verteilungsfunktionen können sofort nach dem Start von **R** mit den folgenden Befehlen ausgewertet werden:

Verteilung von X	Parameter	F_X an Stelle `x` mit **R**
$B(n,p)$	`size`$=n$, `prob`$=p$	`pbinom(x,size,prob)`
Geom(p)	`prob`$=p$	`pgeom(x,prob)`
Pois(λ)	`lambda`$=\lambda$	`ppois(x,lambda)`
Unif(a,b)	`min`$=a$, `max`$=b$	`punif(x,min,max)`
$N(\mu,\sigma^2)$	`mean`$=\mu$, `sd`$=\sqrt{\sigma^2}$	`pnorm(x,mean,sd)`
Exp(λ)	`rate`$=\lambda$	`pexp(x,rate)`

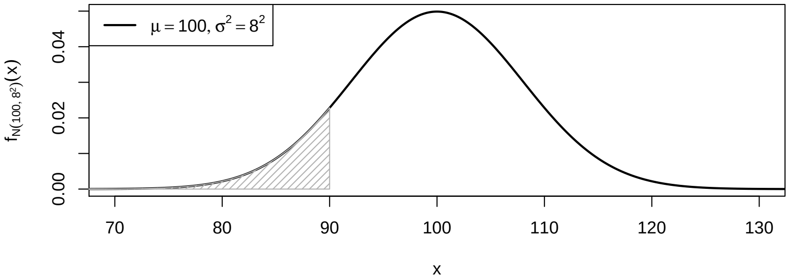
- Ersetzt man in den Befehlen den ersten Buchstaben `p` durch `d` (z.B. `dnorm`), so erhält man den Wert der Dichtefunktion bzw. Wahrscheinlichkeitsfunktion an der Stelle `x`.
- Ersetzt man in den Befehlen den ersten Buchstaben `p` durch `q` (z.B. `qnorm`) und `x` durch `p`, so erhält man das (bzw. ein) `p`-Quantil der zugehörigen Verteilung.
- Ersetzt man schließlich in den Befehlen den ersten Buchstaben `p` durch `r` (z.B. `rnorm`) und `x` durch `n` $\in \mathbb{N}$, so erhält man `n` (Pseudo-)Zufallszahlen zur zugehörigen Verteilung.

Hinweise zur Tabellennutzung

- Bezeichnet $F_{B(n,p)}$ für $n \in \mathbb{N}$ und $p \in (0,1)$ die Verteilungsfunktion der $B(n,p)$-Verteilung, so gilt (!)
$$F_{B(n,1-p)}(x) = 1 - F_{B(n,p)}(n - x - 1)$$
für alle $n \in \mathbb{N}$, $p \in (0,1)$, $x \in \{0, \dots, n-1\}$. Daher werden Tabellen zur Binomialverteilung nur für $p \in (0, 0.5]$ erstellt, und die benötigten Werte für $p \in [0.5, 1)$ mit obiger Formel aus den Werten für $p \in (0, 0.5]$ gewonnen.
- Wegen der Symmetrie der Standardnormalverteilung um 0 gilt nicht nur $\varphi(x) = \varphi(-x)$ für alle $x \in \mathbb{R}$, sondern auch (vgl. Folie 214)
$$\Phi(x) = 1 - \Phi(-x) \qquad \text{für alle } x \in \mathbb{R}\ .$$
Daher werden Tabellen für $\Phi(x)$ in der Regel nur für $x \in \mathbb{R}_+$ erstellt.
- Zur Bestimmung von Quantilen darf in der Klausur ein beliebiger Wert des Intervalls, in dem das Quantil laut Tabelle liegen muss, eingesetzt werden; eine lineare Interpolation ist zwar sinnvoll, aber nicht nötig!
- Generell gilt: Ist ein Wert nicht tabelliert, wird stattdessen ein „naheliegender" Wert aus der Tabelle eingesetzt.
Beispiel: Für fehlenden Wert $F_{B(4,0.28)}(2)$ wird $F_{B(4,0.3)}(2)$ eingesetzt.

Beispiel: Arbeiten mit Normalverteilungstabelle

- **Frage:** Mit welcher Wahrscheinlichkeit nimmt eine $N(100, 8^2)$-verteilte Zufallsvariable Werte kleiner als 90 an? (Wie groß ist die schraffierte Fläche?)

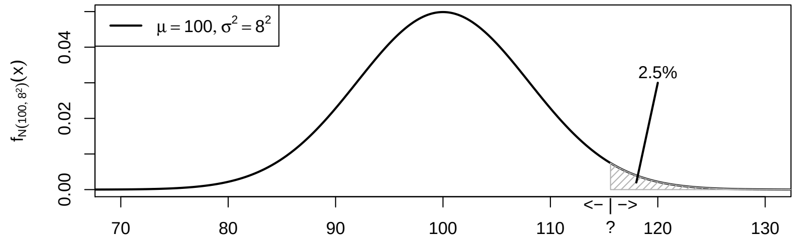


- **Antwort:** Ist $X \sim N(100, 8^2)$, so gilt:

$$\begin{aligned} P\{X < 90\} &= F_{N(100,8^2)}(90) = \Phi\left(\frac{90-100}{8}\right) \\ &= \Phi(-1.25) = 1 - \Phi(1.25) = 1 - 0.8944 = 0.1056 \end{aligned}$$

$\leadsto$ Die gesuchte Wahrscheinlichkeit ist $0.1056 = 10.56\%$.

- **Frage:** Welchen Wert x überschreitet eine $N(100, 8^2)$-verteilte Zufallsvariable nur mit 2.5% Wahrscheinlichkeit? (Welche linke Grenze x führt bei der schraffierten Fläche zu einem Flächeninhalt von 0.025?)



- **Antwort:** Ist $X \sim N(100, 8^2)$, so ist das 97.5%- bzw. 0.975-Quantil von X gesucht. Mit

$$F_X(x) = F_{N(100,8^2)}(x) = \Phi\left(\frac{x-100}{8}\right)$$

erhält man

$$\begin{aligned} \Phi\left(\frac{x-100}{8}\right) \overset{!}{=} 0.975 \quad &\Leftrightarrow \quad \frac{x-100}{8} = \Phi^{-1}(0.975) = 1.96 \\ &\Rightarrow \quad x = 8 \cdot 1.96 + 100 = 115.68 \end{aligned}$$

Beispiel: Arbeiten mit Statistik-Software **R**

- Beantwortung der Fragen (noch) einfacher mit Statistik-Software **R**:
- **Frage:** Mit welcher Wahrscheinlichkeit nimmt eine $N(100, 8^2)$-verteilte Zufallsvariable Werte kleiner als 90 an?
- **Antwort:**

```
> pnorm(90,mean=100,sd=8)
[1] 0.1056498
```

- **Frage:** Welchen Wert x überschreitet eine $N(100, 8^2)$-verteilte Zufallsvariable nur mit 2.5% Wahrscheinlichkeit?
- **Antwort:**

```
> qnorm(0.975,mean=100,sd=8)
[1] 115.6797
```

oder alternativ

```
> qnorm(0.025,mean=100,sd=8,lower.tail=FALSE)
[1] 115.6797
```