Verteilungen

– Bernoulli-Verteilung	Binäre Zufallsgrößen
– Binomialverteilung	Mehrmaliges Durchführen des gleichen Zufallsexperiments
– Stetige Gleichverteilung	Konstante Dichtefunktion in einem Intervall
– Normalverteilung – Standardnormalverteilung	Symmetrische Verteilung um Erwartungswert
– χ^2-Verteilung	Verteilung der Summe von quadrierten Zufallsgrößen

Verteilungen

Bernoulli-Verteilung

- Binäre Zufallsgrößen („0“ oder „1“)
- Verteilung durch $P(X = 1)$ vollständig charakterisiert: $X \backsim Be(p)$

- Es gilt:

$$P(X = 1) = p \; ; \; P(X = 0) = 1-p$$

$$E(X)=p \; ; \; Var(X) = p \cdot (1 - p)$$

Verteilungen

Binomial-Verteilung

– Mehrfache Durchführung des gleichen Zufallsexperiments
– Wie oft tritt Ergebnis A bei n-maligem Durchführen ein
– Bestimmt durch: $X \sim B(n;p)$

Es gilt:

$$P(X = x) = \binom{n}{x} \cdot p^x \cdot q^{n-x} \text{ für } x=0, 1, \ldots, n$$

$$E(X) = n \cdot p$$

$$Var(X) = n \cdot p \cdot (1-p)$$

Verteilungen

Normalverteilung

- Dichtefunktion kann auch als Glockenkurve bezeichnet werden
- Stetige Zufallsgröße, charakterisiert durch Erwartungswert μ und Varianz σ^2: $X \sim N(\mu; \sigma^2)$

- Grafische Darstellung:
  - Glockenkurve, symmetrisch um den Erwartungswert
  - Je größer der Wert σ^2, desto breiter sind die Daten verteilt

- Berechnung von Wahrscheinlichkeiten:

$$Z = \frac{X - \mu}{\sigma}$$

$P(Z \leq 0{,}8416) = F(0{,}8416) = 0{,}8$

$P(-1{,}96 \leq Z \leq 1{,}96) = F_Z(1{,}96) - F_Z(-1{,}96) = 0{,}975 - 0{,}025 = 0{,}95$

Ausgewählte Quantile	
$\Phi(z)$	z
0,9995	3,2905
0,999	3,0902
0,995	2,5758
0,99	2,3263
0,975	1,9600
0,95	1,6449
0,9	1,2816
0,8	0,8416
0,7	0,5244
0,6	0,2533
0,5	0,0000

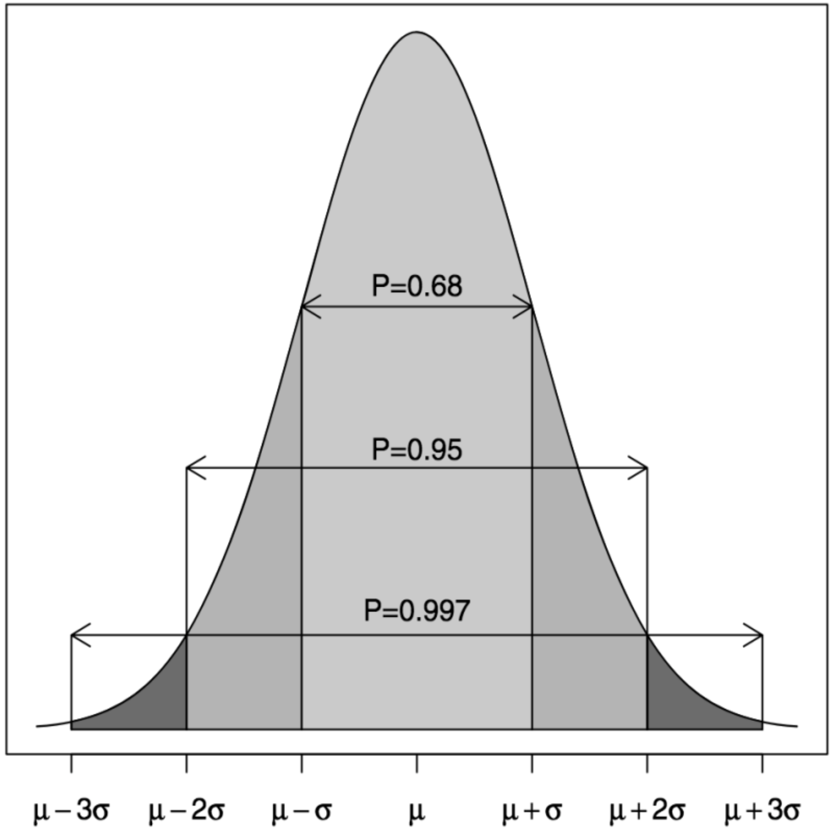
Verteilungen

- Wichtige Wahrscheinlichkeiten der Normalverteilung
  - Wahrscheinlichkeiten im Streuungsbereich $k\sigma$ um den Erwartungswert

$P(\mu - 1\boldsymbol{\sigma} \leq X \leq \mu + 1\boldsymbol{\sigma}) = 0{,}68$

$P(\mu - 2\boldsymbol{\sigma} \leq X \leq \mu + 2\boldsymbol{\sigma}) = 0{,}95$

$P(\mu - 3\boldsymbol{\sigma} \leq X \leq \mu + 3\boldsymbol{\sigma}) = 0{,}997$



Verteilungen

Die χ^2- Verteilung

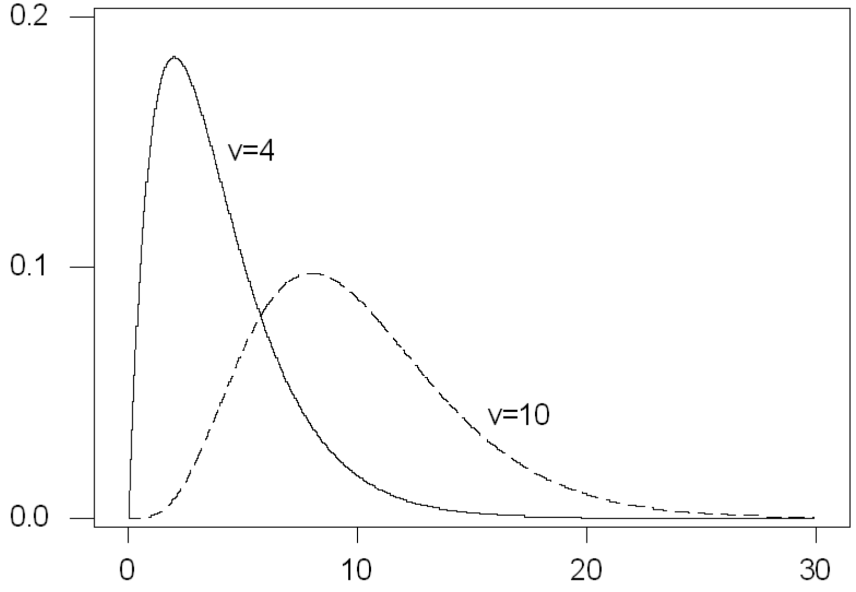
- Wichtig für Schätzen und Testen (Buchkapitel 8)
- Verteilung der Summe von quadrierten normalverteilten Zufallsgrößen
- Charakterisiert durch Parameter „Zahl der Freiheitsgrade" n

$Y \sim \chi^2(n) \qquad X_i \sim N(0\,;1)$

$\Rightarrow Y = \sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$

$E(Y) = n$

$Var(Y) = 2n$



Grenzwertsätze

Gesetz der Großen Zahlen

Ist μ der Erwartungswert der Zufallsgröße X, so liegt der Stichprobenmittelwert $\bar{X}$ für großes n mit hoher Wahrscheinlichkeit sehr nahe bei μ.

$\lim_{n \to \infty} P(|\bar{X} - \mu| \leq \epsilon) = 1$ für beliebig kleines positives ϵ

- Zentraler Grenzwertsatz
  - Ermöglicht Aussage über Verteilung von $\bar{X}$
  - Für große n ist $\bar{X}$ annähernd $N(\mu\,;\,\frac{\sigma^2}{n})$ - verteilt. Faustregel: $n \geq 30$

Vollerhebung vs Teilerhebung

- **Vollerhebung** (Erhebung der Grundgesamtheit)
- Auswahlverfahren
- Stichprobe
- **Teilerhebung**
  - Nicht zufällige Auswahl
    - Willkürliche Auswahl
    - Schneeballverfahren
    - Quotenauswahl
    - Bewusste Auswahl
  - Zufällige Auswahl
    - Gewichtete Auswahl
    - Uneingeschränkte Auswahl
    - Mehrstufige Auswahl

Fehlerarten bei Stichproben

Fehlerarten bei der Erhebung einer Stichprobe

- Bias (Auswahlbias, systematischer Fehler)
- Auswahlfehler (Zufallsfehler)

- Fehlende Werte
- Unvollkommenheit der Auswahlgrundlage
  - Mangelnde Definition der Grundgesamtheit
- Antwortfehler
  - Erinnerungsfehler
  - Soziale Erwünschtheit
  - Fragestellung/ -reihenfolge
- Verarbeitungsfehler

!

Repräsentativität

- Auswahlverfahren vs. Umfang

Uebungs aufgaben

Aufgabe 1.10 der AS

Zu welcher Verteilung gehören die folgenden Zufallsgrößen?

a) X: Messfehler bei der Messung von Körpergrößen
~~b) X: Anzahl von Frauen bei zehn zufällig ausgewählten Teilnehmenden einer Statistik-Vorlesung~~
c) Indikatorvariable für das Ereignis „Über 80% der Teilnehmenden einer Klausur erhalten mehr als 50% der möglichen Punkte.“
d) X: Anzahl der Hörenden einer Statistik-Vorlesung, die den Unterschied zwischen einer Binomial- und einer Hypergeometrischen Verteilung verstanden haben

Aufgabe 1.10 der AS - Lösung

- Normalverteilung
- ~~Hypergeometrische Verteilung~~
- Bernoulliverteilung
- Binomialverteilung

Wieso diese Verteilungen?

Körpergröße und Messfehler typisch für NV

Indikatorvariable: 0 oder 1 als Ausprägung

Unterschied verstanden: Ja/Nein < Bernoulli Experiment wird aber öfter durchgeführt (alle Leute in der StatistikVL)

Aufgabe 1

In einem Biergarten werden die Tischabstände zueinander gemessen. Laut Regierungsverordnung müssen diese 1,5 Meter auseinanderstehen. Genau gemessen weichen die Abstände im Mittel um 6 Zentimeter ab ($\sigma^2 = 0{,}0036$)

- Berechne die Wahrscheinlichkeit, dass ein Tisch weniger als 1,6176 Meter vom nächsten entfernt ist.
- Berechne die Wahrscheinlichkeit, dass ein Tisch weniger als 1,4495 Meter vom nächsten entfernt ist.
- Berechne $P(-1{,}2816 \leq Z \leq 1{,}2816)$.
- In einem anderen Biergarten daneben weichen die Abstände im Durchschnitt um 12 Zentimeter ab. Bei welchem Biergarten wäre die Glockenkurve breiter?

$$Z = \frac{X - \mu}{\sigma}$$

Ausgewählte Quantile

$\Phi(z)$	z
0,9995	3,2905
0,999	3,0902
0,995	2,5758
0,99	2,3263
0,975	1,9600
0,95	1,6449
0,9	1,2816
0,8	0,8416
0,7	0,5244
0,6	0,2533
0,5	0,0000

Aufgabe 1 - Lösung

– 1. Teilaufgabe

$$Z = \frac{1{,}6176 - 1{,}5}{0{,}06} = 1{,}96 \rightarrow P(Z \leq 1{,}96) = F(1{,}96) = 0{,}975$$

– 2. Teilaufgabe

$$Z = \frac{1{,}4495 - 1{,}5}{0{,}06} = -0{,}8416 \rightarrow P(Z \leq -0{,}8416) = 1 - F(0{,}8416) = 1 - 0{,}8 = 0{,}2$$

– 3. Teilaufgabe

$$P(-1{,}2816 \leq Z \leq 1{,}2816) = F(1{,}2816) - F(-1{,}2816) = 0{,}9 - (1-0{,}9) = 0{,}9 - 0{,}1 = 0{,}8$$

– 4. Teilaufgabe

$\sigma^2 = 0{,}12^2 = 0{,}0144 > 0{,}0036$ -> Die Glockenkurve des Nachbar-Biergartens wäre breiter.

Aufgabe 1.7 der AS

In einer Urne befinden sich vier weiße und sechs schwarze Kugeln.

a) Geben sie die Wahrscheinlichkeitsfunktion der Zufallsvariable X „Anzahl der weißen Kugeln in einer Stichprobe vom Umfang $n = 1$" an und berechnen Sie den Erwartungswert $E(X)$ und die Varianz $Var(X)$.
b) Wie ist die Zufallsvariable Y „Anzahl der weißen Kugeln in einer Stichprobe vom Umfang $n = 4$" verteilt, wenn man die Kugeln mit Zurücklegen entnimmt? Geben Sie jeweils die zugehörige Wahrscheinlichkeitsverteilung an und ermitteln sie jeweils $E(Y)$ und $Var(Y)$.

Aufgabe 1.7 der AS - Lösung

a)

X ist eine bernoulli-verteilte Zufallsgröße, da sie nur die Werte 0 und 1 annehmen kann.

x	f(x)
0	0,6
1	0,4

- $E(X) = p = 0{,}4$
- $Var(X) = p \cdot (1-p) = 0{,}4 \cdot 0{,}6 = 0{,}24$

b)

Y ist eine binomialverteilte Zufallsgröße mit $n = 4$ und $p = 0{,}4$: $Y \sim B(4;0{,}4)$.

$P(Y = y) = \binom{n}{y} \cdot p^y \cdot q^{n-y}$

y	P(Y = y)	f(y)
0	$P(Y = 0) = \binom{4}{0} \cdot 0{,}4^0 \cdot 0{,}6^{4-0}$	0,1296
1	$P(Y = 1) = \binom{4}{1} \cdot 0{,}4^1 \cdot 0{,}6^{4-1}$	0,3456
2	$P(Y = 2) = \binom{4}{2} \cdot 0{,}4^2 \cdot 0{,}6^{4-2}$	0,3456
3	$P(Y = 3) = \binom{4}{3} \cdot 0{,}4^3 \cdot 0{,}6^{4-3}$	0,1536
4	$P(Y = 4) = \binom{4}{4} \cdot 0{,}4^4 \cdot 0{,}6^{4-4}$	0,0256

- $E(Y) = n \cdot p = 4 \cdot 0{,}4 = 1{,}6$
- $Var(Y) = n \cdot p \cdot (1-p) = 4 \cdot 0{,}4 \cdot 0{,}6 = 0{,}96$

Aufgabe 1.8 der AS

Eine Bäckerei setzt zur Produktion ihrer Brezeln zwei unabhängig voneinander arbeitende Maschinen A und B ein. Pro Tag werden auf Maschine A 870, auf Maschine B 800 Brezeln produziert, wobei der Ausschussanteil bei Maschine A erfahrungsgemäß bei 10% und bei Maschine B 5% beträgt.

a) Wie sind die Zufallsvariablen X_A und X_B „Tagesproduktion an einwandfreien Brezeln auf Maschine A bzw. B“ verteilt?

b) Geben Sie den Erwartungswert und die Varianz der Zufallsvariablen X_A und X_B an.

Aufgabe 1.8 der AS - Lösung

a)

$X_A \sim B(870;0,9)$ $X_B \sim B(800;0,95)$

b)

- $E(X_A) = n \cdot p = 870 \cdot 0,9 = 783$
  $E(X_B) = n \cdot p = 800 \cdot 0,96 = 760$

- $Var(X_A) = n \cdot p \cdot (1\text{-}p) = 870 \cdot 0,9 \cdot 0,1 = 78,3$
  $Var\ (X_B) = n \cdot p \cdot (1\text{-}p) = 800 \cdot 0,95 \cdot 0,05 = 38$

it`s time for a quiz!