# 6. Vorlesung Statistik I
# **Einfache Zufallsstichprobe, Punktschätzung**

We are happy to share our materials openly:
The content of these Open Educational Resources by Lehrstuhl für Psychologische Methodenlehre und Diagnostik, Ludwig-Maximilians-Universität München is licensed under CC BY-SA 4.0. The CC Attribution-ShareAlike 4.0 International license means that you can reuse or transform the content of our materials for any purpose as long as you cite our original materials and share your derivatives under the same license.

# Populationen, Stichproben und Wahrscheinlichkeitstheorie

# Wiederholung I



- Die **Inferenzstatistik** umfasst alle statistischen Verfahren, die es erlauben, trotz der Informationsunvollständigkeit der Stichprobendaten Aussagen über eine Population zu treffen.

- Definition **Population:** Gesamtheit aller Merkmalsträger*innen, auf die eine Untersuchungsfrage gerichtet ist.

- Definition **Stichprobe:** Auswahl bestimmter Merkmalsträger*innen aus einer Population

# Wiederholung II



- Problem:
    - Wenn nur ein Teil der Grundgesamtheit erfasst wird, z.B. 500 Personen, ist die Informationslage in Bezug auf die Untersuchungsfrage unvollständig. Wir können nicht einfach deskriptiv-statistische Methoden verwenden.
    - Wie kann man trotzdem Aussagen treffen, die sich auf alle Personen der Grundgesamtheit beziehen, obwohl nur die Daten einer Stichprobe vorliegen?
- Idee:
    - Wir ziehen die Personen **zufällig** aus der Population in die Stichprobe.
    - Wir greifen auf mathematische Methoden zur Formalisierung von Zufallsprozessen zurück → **Wahrscheinlichkeitstheorie**
    - Aus diesen ergeben sich Methoden, die Rückschlüsse von der Stichprobe auf die Population erlauben → **Inferenzstatistik**

# Beispiele



- Da praktisch alle psychologischen Theorien Aussagen über Populationen enthalten, sind zu ihrer empirischen Überprüfung **immer** inferenzstatistische Methoden notwendig.

- Beispiele für psychologische Fragestellungen:
  - Beispiel I:
    - Wir interessieren uns für die relative Häufigkeit $h_D$ der Personen in Deutschland, die an einer Depression erkrankt sind.
  - Beispiel II:
    - Wir interessieren uns für den Mittelwert $\bar{x}_{IQ}$ und die empirische Varianz $s^2_{emp\ IQ}$ des Intelligenzquotienten (IQ) von Student*innen in Deutschland.

# Offene Fragen



- Idee:
  - Wir ziehen die Personen **zufällig** aus der Population in die Stichprobe.
  - Wir greifen auf mathematische Methoden zur Formalisierung von Zufallsprozessen zurück → **Wahrscheinlichkeitstheorie**
  - Aus diesen ergeben sich Methoden, die Rückschlüsse von der Stichprobe auf die Population erlauben → **Inferenzstatistik**

- Offene Fragen:
  - Wie lässt sich das zufällige Ziehen von Personen mithilfe der Wahrscheinlichkeitstheorie formalisieren?
  - Wie lassen sich hieraus inferenzstatistische Methoden ableiten?

# Offene Fragen



- Offene Fragen:
  - **Wie lässt sich das zufällige Ziehen von Personen mithilfe der Wahrscheinlichkeitstheorie formalisieren?**
  - Wie lassen sich hieraus inferenzstatistische Methoden ableiten?

# Zufällige Ziehung einer einzelnen Person



- Zunächst: Zufällige Ziehung einer einzelnen Person aus einer Population von *N* Personen.

- Dieser Vorgang ist ein Zufallsexperiment:
  - Wir wissen im Voraus nicht, welche Person gezogen wird.
  - Die Ergebnismenge $\Omega$ ist die Menge aller Personen in der Population:

$$\Omega = \{Person\ 1, Person\ 2, \dots, Person\ i, \dots, Person\ N\}$$

- Wir setzen voraus, dass jede Person *i* in der Population die gleiche Wahrscheinlichkeit hat, gezogen zu werden. Das heißt, alle Elementarereignisse haben die gleiche Wahrscheinlichkeit:

$$P(\{Person\ i\}) = \frac{1}{N}$$

# Beispiel Depression I



- Wir interessieren uns für die relative Häufigkeit $h_D$ der Personen in Deutschland, die an einer Depression erkrankt sind.

- Sei $N_D$ die Anzahl der depressiven Personen in der Population und $A_D$ die Menge der depressiven Personen in der Population:

$$A_D = \{Depr.Person\ 1, Depr.Person\ 2, \dots, Depr.Person\ N_D\}$$

- Die relative Häufigkeit der Depression in der Population ist also $h_D = \frac{N_D}{N}$

- Die Wahrscheinlichkeit, zufällig eine depressive Person zu ziehen ist:

$$P(A_D) = P(\{Depr.Person\ 1\}) + P(\{Depr.Person\ 2\}) + \cdots + P(\{Depr.Person\ N_D\})$$

$$= \frac{1}{N} + \cdots + \frac{1}{N} = \frac{N_D}{N} = h_D$$

# Beispiel Depression II



- Die Wahrscheinlichkeit dafür, zufällig eine depressive Person zu ziehen, entspricht also der relativen Häufigkeit der Depression in der Population:

$$P(A_D) = h_D$$

- Intuitiv: Die Population entspricht einer „Urne“, die Personen den „Kugeln“ in der Urne und die Wahrscheinlichkeit $P(A_D)$ ist die Wahrscheinlichkeit dafür, eine depressive „Kugel“ zu ziehen.

# Beispiel Depression III



- Sei nun *X* eine Zufallsvariable, die den Wert 1 annimmt, falls die zufällig gezogene Person depressiv ist, und 0, falls sie es nicht ist.

- Diese Zufallsvariable ist eine Bernoulli-Variable und folgt somit einer Bernoulli-Verteilung.

- Der Parameter $\pi$ der Bernoulli-Verteilung entspricht der Wahrscheinlichkeit, dass *X* den Wert 1 annimmt, also der Wahrscheinlichkeit, eine depressive Person zu ziehen. Diese Wahrscheinlichkeit entspricht wiederum der relativen Häufigkeit der Depression in der Population (siehe letzte Folie).

- Formal:

$$\pi = P(X = 1) = \ldots = P(A_D) = h_D$$

# Beispiel Depression IV



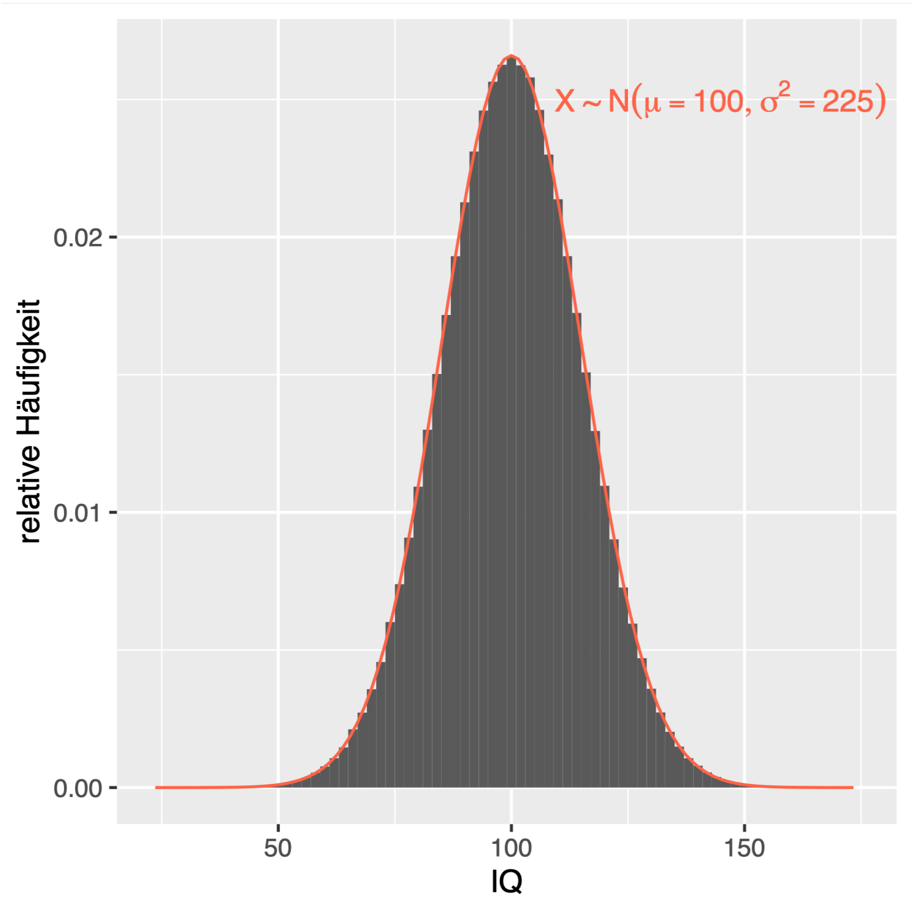
- Zusammengefasst: Unter der Voraussetzung, dass
  - jede Person in der Population die gleiche Wahrscheinlichkeit hat, gezogen zu werden,
  - *X* eine Zufallsvariable ist, die den Wert 1 annimmt, falls die gezogene Person depressiv ist, und 0, falls nicht,

  **folgt *X* einer Bernoulli-Verteilung** und der **Wert des Parameters $\pi$** dieser Bernoulli-Verteilung ist **identisch mit dem Wert der relativen Häufigkeit $h_D$ der Depression in der Population**.

- Wenn wir herausfinden wollen, wie hoch die relative Häufigkeit der Depression in der Population ist, müssen wir lediglich herausfinden, welchen Wert der Parameter $\pi$ hat.
- Wenn wir z.B. wüssten, dass $\pi = 0.3$ ist, wüssten wir auch, dass die relative Häufigkeit der Depression in der Population $h_D = 0.3$ ist.
- Da $\pi$ ein Parameter einer Wahrscheinlichkeitsverteilung ist, können wir das Problem der Bestimmung einer deskriptivstatistischen Maßzahl in der Population ($h_D$) komplett in die Wahrscheinlichkeitstheorie verlagern und somit alle Mittel verwenden, die uns diese zur Verfügung stellt.

# Beispiel Intelligenz I






- Wir interessieren uns für den Mittelwert $\bar{x}_{IQ}$ und die empirische Varianz $s^2_{emp\ IQ}$ des IQs von Personen in Deutschland.
- Wir setzen voraus, dass das Histogramm der Variable IQ in der Population der Personen in Deutschland durch die Wahrscheinlichkeitsdichtefunktion einer Normalverteilung approximiert werden kann, d.h. dass das Histogramm die „Form" der Dichte einer Normalverteilung hat.
- Dies ist eine **Annahme**, von der wir nicht wissen, ob sie zutrifft. Wir werden jedoch Methoden kennenlernen, um die Plausibilität dieser Annahme zu überprüfen.

# Beispiel Intelligenz II



- Außerdem setzen wir wieder voraus, dass alle Personen die gleiche Wahrscheinlichkeit haben, gezogen zu werden.

- Sei nun *X* eine Zufallsvariable, die für den IQ der zufällig gezogenen Person steht.

- Man kann dann beweisen, dass diese Zufallsvariable *X* einer Normalverteilung folgt und

  - der Parameter $\mu$ dieser Normalverteilung dem Mittelwert des IQs in der Population entspricht:

$$\mu = \bar{x}_{IQ}$$

  - der Parameter $\sigma^2$ dieser Normalverteilung der empirischen Varianz des IQs in der Population entspricht:

$$\sigma^2 = s^2_{emp\ IQ}$$

- Der Beweis hierfür funktioniert ähnlich wie bei der Bernoulli-Verteilung, ist aber deutlich aufwendiger.

# Beispiel Intelligenz III

- Zusammengefasst: Unter der Voraussetzung, dass
  - Das Histogramm des IQs in der Population der Personen in Deutschland durch die Wahrscheinlichkeitsdichtefunktion einer Normalverteilung approximiert werden kann,
  - jede Person in der Population die gleiche Wahrscheinlichkeit hat, gezogen zu werden,
  - *X* eine Zufallsvariable ist, die für den IQ der gezogenen Person steht,

**folgt *X* einer Normalverteilung** und

- der **Wert des Parameters $\boldsymbol{\mu}$** dieser Normalverteilung ist **identisch mit dem Mittelwert $\boldsymbol{\bar{x}_{IQ}}$ des IQs in der Population**,
- der **Wert des Parameters $\boldsymbol{\sigma^2}$** dieser Normalverteilung ist **identisch mit der empirischen Varianz $\boldsymbol{s^2_{emp\ IQ}}$ des IQs in der Population**.

# Beispiel Intelligenz IV



- Wenn wir herausfinden wollen, was der mittlere IQ in der Population ist, müssen wir lediglich herausfinden, welchen Wert der Parameter $\mu$ hat.

- Wenn wir herausfinden wollen, was die empirische Varianz des IQs in der Population ist, müssen wir lediglich herausfinden, welchen Wert der Parameter $\sigma^2$ hat.

- Auch in diesem Fall können wir also das Problem der Bestimmung von deskriptivstatistischen Maßzahlen in der Population ($\bar{x}_{IQ}$ und $s^2_{emp\ IQ}$) komplett in die Wahrscheinlichkeitstheorie verlagern und somit alle Mittel verwenden, die uns diese zur Verfügung stellt.

# Einfache Zufallsstichprobe



- Jetzt: Zufällige Ziehung von $n$ Personen aus einer Population von $N$ Personen (mit „Zurücklegen“).
- $n$ wird hierbei **Stichprobenumfang** genannt.
- Dieser Vorgang ist ein Zufallsexperiment:
  - Wir wissen im Voraus nicht, welche Personen gezogen werden
  - Die Ergebnismenge $\Omega$ ist die Menge aller möglichen Stichproben mit $n$ Personen.
- Wir setzen voraus, dass jede Person in der Population bei jeder der $n$ Ziehungen die **gleiche Wahrscheinlichkeit** hat, in die Stichprobe gezogen zu werden.
- Wir setzen zudem voraus, dass die $n$ Personen **unabhängig** voneinander gezogen werden.
- Zufallsstichproben mit diesen beiden Eigenschaften nennt man **einfache Zufallsstichproben**.

# iid Zufallsvariablen I



- Seien nun $X_1$, $X_2$, …, $X_i$, …, $X_n$ Zufallsvariablen, die für die Werte der Personen in der Stichprobe auf einer uns interessierenden Variable stehen. In unseren Beispielen:
  - *Beispiel 1*: $X_i$ nimmt den Wert 1 an, falls die *i*-te zufällig gezogene Person depressiv ist, und 0 falls nicht
  - *Beispiel 2*: $X_i$ ist der IQ der *i*-ten zufällig gezogenen Person

- Unter der Voraussetzung einer einfachen Zufallsstichprobe haben alle diese Zufallsvariablen die gleiche Wahrscheinlichkeitsverteilung, in unseren Beispielen:
  - *Beispiel 1*: $X_i \sim Be(\pi)$ für alle *i* = 1, 2, …, *n*
  - *Beispiel 2*: $X_i \sim N(\mu, \sigma^2)$ für alle *i* = 1, 2, …, *n*

- Wie im Fall der Ziehung einer einzelnen Person, entsprechen die Parameter $\pi$, $\mu$ und $\sigma^2$ wieder den uns jeweils interessierenden deskriptivstatistischen Maßzahlen in der Population ($\pi = h_D$ im 1. Beispiel, $\mu = \bar{x}_{IQ}$ und $\sigma^2 = s^2_{emp\ IQ}$ im 2. Beispiel).

- Außerdem sind die Zufallsvariablen $X_1$, $X_2$, …, $X_n$ bei unabhängiger Ziehung der *n* Personen unabhängig.

# iid Zufallsvariablen II



- Unabhängige Zufallsvariablen $X_1$, $X_2$, …, $X_n$, die alle die gleiche Wahrscheinlichkeitsverteilung aufweisen, nennen wir **iid Zufallsvariablen**.

- iid steht hierbei für „**i**ndependent and **i**dentically **d**istributed“

- Notation:

$$X_i \overset{\text{iid}}{\sim} Be(\pi)$$

$$X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$$

- Das Vorliegen von iid Zufallsvariablen ist eine Voraussetzung für die meisten inferenzstatistischen Verfahren.

# Zusammenfassung I



- Falls wir uns für die **relative Häufigkeit** einer Messwertausprägung einer **diskreten** Variable in der Population interessieren:
  - Wir ziehen eine einfache Zufallsstichprobe der Größe *n* aus der Population.
  - Wir betrachten die Zufallsvariablen $X_1$, $X_2$, …, $X_i$, …, $X_n$ , wobei $X_i$ den Wert 1 annimmt, falls die *i*-te zufällig gezogene Person die uns interessierende Messwertausprägung aufweist, und 0 falls nicht.
  - Für diese Zufallsvariablen gilt dann:
    - $X_i \overset{\text{iid}}{\sim} Be(\pi)$
    - der Parameter $\pi$ entspricht der relativen Häufigkeit der uns interessierenden Messwertausprägung in der Population.

# Zusammenfassung II



- Falls wir uns für den **Mittelwert** und die **empirische Varianz** einer **stetigen** Variable in der Population interessieren, deren Histogramm durch die Wahrscheinlichkeitsdichtefunktion einer Normalverteilung approximiert werden kann:
  - Wir ziehen eine einfache Zufallsstichprobe der Größe *n* aus der Population.
  - Wir betrachten die Zufallsvariablen $X_1, X_2, \ldots, X_i, \ldots, X_n$ , wobei $X_i$ für den Wert der *i*-ten zufällig gezogenen Person auf der uns interessierenden stetigen Variable steht.
  - Für diese Zufallsvariablen gilt dann:
    - $X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$
    - der Parameter $\mu$ entspricht dem Mittelwert der stetigen Variable in der Population.
    - der Parameter $\sigma^2$ entspricht der empirischen Varianz der stetigen Variable in der Population.

# Probleme



- Leider werden in der Praxis oft keine einfachen Zufallsstichproben gezogen. Häufig auftretende Probleme:
- **Fehlende Repräsentativität**:
  - Eine bestimmte Teilgruppe von Personen in der Population hat eine höhere Wahrscheinlichkeit, in die Stichprobe gezogen zu werden, als andere Personen.
  - Beispiel: Nur Personen die Psychologie studieren, haben eine echt positive Wahrscheinlichkeit, gezogen werden.
  - Folge: Die interessierende Maßzahl in der Population entspricht nicht mehr dem Parameter der jeweiligen Wahrscheinlichkeitsverteilung. Inferenzstatistische Verfahren, die hierauf aufbauen, sind verzerrt.
- **Abhängigkeit der Ziehungen**:
  - Die Ziehungen der Personen sind nicht unabhängig voneinander.
  - Zum Beispiel sog. geschachtelte Stichproben: Zufälliges Ziehen einer Schule, dann zufälliges Ziehen von Personen aus dieser Schule.
  - Folge: Die Zufallsvariablen sind nicht mehr iid. Inferenzstatistische Verfahren, die hierauf aufbauen, sind verzerrt.
  - Falls die Art der Abhängigkeit bekannt ist, kann dies mithilfe statistischer Methoden ausgeglichen werden.

# Offene Fragen

- Offene Fragen:
  - Wie lässt sich das zufällige Ziehen von Personen mithilfe der Wahrscheinlichkeitstheorie formalisieren?
  - **Wie lassen sich hieraus inferenzstatistische Methoden ableiten?**

# Überblick Inferenzstatistik

# Inferenzstatistik



- Wir haben gesehen, dass sich Aussagen über Populationen im Fall von einfachen Zufallsstichproben auf Aussagen über Parameter von Wahrscheinlichkeitsverteilungen reduzieren lassen.
- Aber: Dies hat das Problem des Rückschlusses von Stichproben auf Populationen erst einmal nur in die Wahrscheinlichkeitstheorie verlagert.
- Es ist immer noch unklar, wie man von Stichproben zu Aussagen über Parameter von Wahrscheinlichkeitsverteilungen gelangt.
- Dies ist Gegenstand der **Inferenzstatistik**.

# Parameterschätzung und Statistische Hypothesentests



- Je nach Fragestellung, d.h. je nach Art der Aussage, die über einen Parameter getroffen werden soll, lassen sich inferenzstatistische Methoden zwei Gebieten zuordnen:

- **Parameterschätzung**:
  - Fragestellung: Welchen konkreten Wert hat ein Parameter?
  - Beispiele:
    - Welchen Wert hat $\pi$?
    - Welche Werte haben $\mu$ und $\sigma^2$?
  - Resultate von Parameterschätzungen sind konkrete Zahlen oder Intervalle.

- **Statistische Hypothesentests**:
  - Fragestellung: Entspricht ein Parameter einem vorgegebenen Wert oder liegt er in einem vorgegebenen Bereich?
  - Beispiele:
    - Ist $\pi = 0.5$?
    - Ist $\mu = 100$?
    - Ist $\sigma^2 > 10$?
  - Das Resultat eines statistischen Hypothesentests ist eine Ja/Nein-Antwort.

# Ausblick



- Wir werden uns in den nächsten Sitzungen mit Methoden der Parameterschätzung befassen.

- Im Anschluss daran werden wir uns für den Rest des Semesters mit statistischen Hypothesentests beschäftigen.

# Parameterschätzung

# Punktschätzung und Intervallschätzung I



- Ziel der Parameterschätzung: Auf der Basis einer einfachen Zufallsstichprobe eine Aussage darüber treffen, welchen konkreten Wert ein Parameter hat.

- Man unterscheidet in der Parameterschätzung zwischen
  - Punktschätzung und
  - Intervallschätzung

- **Punktschätzung**:
  - Ergebnis einer Punktschätzung ist eine konkrete Zahl.
  - Beispiel: Wir gehen auf Basis unserer Stichprobe davon aus, dass der Parameter $\pi$ gleich 0.45 ist.

- **Intervallschätzung**
  - Ergebnis einer Intervallschätzung ist ein Intervall von Zahlen.
  - Beispiel: Wir gehen auf Basis unserer Stichprobe davon aus, dass der Parameter $\mu$ im Bereich zwischen 101 und 108 liegt.

# Punktschätzung und Intervallschätzung II



- Zur Erinnerung (sehr wichtig!): Aussagen über Parameter sind im Fall von einfachen Zufallsstichproben gleichbedeutend mit Aussagen über die uns eigentlich interessierenden deskriptivstatistischen Maßzahlen in der Population.
  - „Wir gehen auf Basis unserer Stichprobe davon aus, dass der Parameter $\pi$ gleich 0.45 ist." ist gleichbedeutend mit „Wir gehen auf Basis unserer Stichprobe davon aus, dass die relative Häufigkeit der uns interessierenden Messwertausprägung in der Population gleich 0.45 ist."
  - „Wir gehen auf Basis unserer Stichprobe davon aus, dass der Parameter $\mu$ im Bereich zwischen 101 und 108 liegt." ist gleichbedeutend mit „Wir gehen auf Basis unserer Stichprobe davon aus, dass der Mittelwert der uns interessierenden stetigen Variable in der Population im Bereich zwischen 101 und 108 liegt."

# Punktschätzung

# Schätzwerte I



- Wir interessieren uns für eine deskriptivstatistische Maßzahl in einer Population.
  - *Beispiel 1*:
    - die relative Häufigkeit $h_D$ der Depression in der Population
  - *Beispiel 2*:
    - den Mittelwert $\bar{x}_{IQ}$ des IQs in der Population
    - die empirische Varianz $s^2_{emp\ IQ}$ des IQs in der Population
- Wir ziehen eine einfache Zufallsstichprobe und betrachten die iid Zufallsvariablen $X_1$, $X_2$, …, $X_n$ mit ihren Realisationen $x_1$, $x_2$, …, $x_n$ (dies sind die Messwerte der Personen in der Stichprobe).
- Diese ZVs weisen eine Wahrscheinlichkeitsverteilung auf, deren Parameter jeweils den uns interessierenden deskriptivstatistischen Maßzahlen in der Population entsprechen:
  - *Beispiel 1*: $\pi = h_D$
  - *Beispiel 2*: $\mu = \bar{x}_{IQ}$ und $\sigma^2 = s^2_{emp\ IQ}$

# Schätzwerte II



- Wir wollen auf Basis der Realisationen der ZVs für jeden Parameter jeweils einen **Schätzwert** bestimmen, d.h. einen Wert, von dem wir auf der Basis unserer Stichprobe davon ausgehen, dass er „möglichst nahe“ an dem wahren Parameterwert bzw. dem wahren Wert der Maßzahl in der Population liegt.

- Naheliegende Idee: Wir verwenden als Schätzwert einfach zunächst die entsprechende deskriptivstatistische Maßzahl in der Stichprobe.

- Ob diese Schätzungen unseren qualitativen Ansprüchen genügen, muss sich aber noch herausstellen.

# Schätzwerte III



- Die relative Häufigkeit der Depression in der Stichprobe als naheliegenden Schätzwert für $\pi$:

$$h(x_2) = h(1) = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

- Den Mittelwert des IQs in der Stichprobe als naheliegenden Schätzwert für $\mu$:

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

- Die empirische Varianz des IQs in der Stichprobe als naheliegenden Schätzwert für $\sigma^2$:

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

# Schätzwerte IV



- Wie können wir feststellen, ob diese Schätzwerte jeweils nahe an den wahren Parametern liegen?
- Antwort: Für eine einzelne Stichprobe können wir dies leider nicht.
- Um zu bestimmen, wie weit ein einzelner Schätzwert vom wahren Parameterwert entfernt ist, müssten wir ja den wahren Parameterwert schon kennen.
- Aber: Da die Ziehung der einfachen Zufallsstichprobe ein Zufallsexperiment ist, können wir uns anschauen, wie sich die Schätzwerte verhalten würden, falls wir wiederholt eine solche einfache Zufallsstichprobe ziehen und die Schätzwerte berechnen würden.
- Wir können auf diese Art und Weise herausfinden, wie gut unsere Schätzungen „im Durchschnitt“ sind.

# Schätzwerte V



- Jedes Mal, wenn wir eine einfache Zufallsstichprobe der Größe *n* aus der Population ziehen, gelangen andere Personen in unsere Stichprobe.

- Daher ergeben sich auch jedes Mal andere Realisationen $x_1$, $x_2$, …, $x_n$ der Zufallsvariablen $X_1$, $X_2$, …, $X_n$ und daher auch andere Schätzwerte.

- Die Schätzwerte unterliegen also Zufallsschwankungen.

# Schätzwerte VI



- Beispiel: Unsere Population besteht aus den folgenden Personen:
  - Peter – depressiv
  - Petra – nicht depressiv
  - Paul – nicht depressiv
  - Paula – depressiv
  - Anna – nicht depressiv

- Der relative Häufigkeit der Depression in dieser Population wäre

$$h_D = \frac{2}{5} = 0.4$$

- Angenommen, wir würden uns für $h_D$ in dieser Population interessieren, könnten die Population aber nicht direkt untersuchen, sondern nur einfache Zufallsstichproben der Größe $n$ = 3 ziehen.

- Zur Erinnerung: In diesem Fall ist $h_D = \pi$

# Schätzwerte VII



- Beispiele für mögliche zufällig gezogene Stichproben der Größe *n* = 3 und jeweils resultierender Schätzwert $\bar{x}$ für $\pi$ bzw. $h_D$ :

  - Stichprobe: Peter, Paula, Anna
    - Messwerte: $x_1 = 1$, $x_2 = 1$, $x_3 = 0$
    - Schätzwert: $\bar{x} \approx 0.67$

  - Stichprobe: Paul, Paula, Anna
    - Messwerte: $x_1 = 0$, $x_2 = 1$, $x_3 = 0$
    - Schätzwert: $\bar{x} \approx 0.33$

  - Stichprobe: Anna, Paul, Petra
    - Messwerte: $x_1 = 0$, $x_2 = 0$, $x_3 = 0$
    - Schätzwert: $\bar{x} = 0$
  - etc.

- Jedes Mal würde aufgrund der zufälligen Ziehung der Personen ein zufälliger anderer Schätzwert resultieren!

# Schätzwerte VIII



- Diese Zufälligkeit der Schätzwerte können wir wieder mithilfe der Wahrscheinlichkeitstheorie beschreiben.

- Hieraus werden sich dann Gütekriterien für die Punktschätzung ergeben.

# Schätzfunktionen I



- Das Ziehen einer einfachen Zufallsstichprobe ist ein Zufallsexperiment.
- Das Ergebnis dieses Zufallsexperiments ist eine Stichprobe von *n* Personen.
- Ein Schätzwert ist eine konkrete Zahl und kann für jede dieser Stichproben aus den Realisationen $x_1, x_2, \ldots, x_n$ der iid Zufallsvariablen $X_1, X_2, \ldots, X_n$ berechnet werden.
- Wir können also eine Zufallsvariable definieren, die jedem möglichen Ergebnis des Zufallsexperiments – also jeder möglichen Stichprobe - den aus ihr berechneten Schätzwert zuweist.
- Eine solche Zufallsvariable heißt **Schätzfunktion**.
- Der konkrete **Schätzwert** ist dann die **Realisation dieser Schätzfunktion**.

# Schätzfunktionen II

- Beispiele:

  - Die zu dem Schätzwert

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

  gehörige Schätzfunktion für den Parameter $\pi$ ist die Zufallsvariable

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

  - Der konkrete Schätzwert $\bar{x}$ ist eine Realisation der Schätzfunktion $\bar{X}$

# Schätzfunktionen III



- Beispiele:

  - Die zu dem Schätzwert

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

gehörige Schätzfunktion für den Parameter $\mu$ ist die Zufallsvariable

$$\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

  - Der konkrete Schätzwert $\bar{x}$ ist eine Realisation der Schätzfunktion $\bar{X}$

# Schätzfunktionen IV

- Beispiele:
  - Die zu dem Schätzwert

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

gehörige Schätzfunktion für den Parameter $\sigma^2$ ist die Zufallsvariable

$$S_{emp}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

  - Der konkrete Schätzwert $s_{emp}^2$ ist eine Realisation der Schätzfunktion $S_{emp}^2$

# Schätzfunktionen V



- Bemerkung I:
  - Schätzwerte sind Funktionen der Realisationen $x_i$ der Zufallsvariablen $X_i$
  - Schätzfunktionen sind Funktionen der Zufallsvariablen $X_i$ selbst

- Bemerkung II: Schätzfunktionen sind Zufallsvariablen, die allein dadurch zu Schätzfunktionen für einen Parameter werden, dass man sie zur Schätzung für diesen Parameter verwendet. **Das heißt aber natürlich nicht, dass es sich damit automatisch um *gute* Schätzfunktionen für diesen Parameter handelt.**

# Notation I



- **Parameter** werden meist mit **griechischen Kleinbuchstaben** bezeichnet:
  - $\mu$
  - $\sigma^2$
  - $\pi$

- **Allgemeine Schätzfunktionen** für diese Parameter werden mit dem entsprechenden **Symbol des Parameters mit Dach** bezeichnet:
  - $\hat{\mu}$ bezeichnet eine Schätzfunktion für den Parameter $\mu$
  - $\hat{\sigma}^2$ bezeichnet eine Schätzfunktion für den Parameter $\sigma^2$
  - $\hat{\pi}$ bezeichnet eine Schätzfunktion für den Parameter $\pi$

- **Konkrete Schätzfunktionen** werden mit den **Großbuchstaben** der Zufallsvariablen bezeichnet, denen sie entsprechen, z.B.:
  - $\bar{X}$
  - $S^2_{emp}$

# Notation II



- **Allgemeine Schätzwerte** für einen Parameter werden mit dem entsprechenden **Symbol des Parameters mit Dach und dem Subskript „Wert"** bezeichnet:
  - $\hat{\mu}_{Wert}$ bezeichnet einen Schätzwert für den Parameter $\mu$
  - $\hat{\sigma}^2_{Wert}$ bezeichnet einen Schätzwert für den Parameter $\sigma^2$
  - $\hat{\pi}_{Wert}$ bezeichnet einen Schätzwert für den Parameter $\pi$

- **Konkrete Schätzwerte** werden mit dem **Kleinbuchstaben** der Zufallsvariablen bezeichnet, deren Realisationen sie sind z.B.:
  - $\bar{x}$
  - $s^2_{emp}$

# Parameter, Schätzfunktionen und Schätzwerte



- Die Unterscheidung zwischen Parametern, Schätzfunktionen und Schätzwerten ist für das Verständnis sehr wichtig.

- Wir werden in der Inferenzstatistik noch sehr häufig der folgenden Situation begegnen:
  - Wir interessieren uns für einen Parameter einer Wahrscheinlichkeitsverteilung.
  - Für diesen Parameter bestimmen wir eine Schätzfunktion.
  - Wir ziehen eine einfache Zufallsstichprobe, in der sich die Schätzfunktion in einem konkreten Schätzwert für den Parameter realisiert.

# Eigenschaften von Schätzfunktionen

- Da Schätzfunktionen Zufallsvariablen sind, haben sie – **wie alle Zufallsvariablen** – eine Wahrscheinlichkeitsverteilung, einen Erwartungswert, eine Varianz und eine Standardabweichung.

- Mit konkreten Wahrscheinlichkeitsverteilungen von Schätzfunktionen werden wir uns in der nächsten Vorlesung beschäftigen.

- Aber schon anhand der Eigenschaften Erwartungswert, Varianz und Standardabweichung lässt sich einschätzen, ob und wie gut sie zur Schätzung eines unbekannten Parameters geeignet ist.

- Für die Schätzung eines beliebigen Parameters lassen sich also für eine Schätzfunktion sogenannte **Gütekriterien** formulieren, die die Verwendung der Funktion als Schätzfunktion rechtfertigt.

# Gütekriterien einer Schätzfunktion

# Gütekriterien für Schätzfunktionen



Im Folgenden sei $\theta$ ein beliebiger Parameter einer beliebigen Wahrscheinlichkeitsverteilung und $\hat{\theta}$ eine Schätzfunktion für diesen Parameter.

- Wir können uns damit nun erst mal unabhängig von einer konkreten Schätzfunktion oder eines Parameters überlegen, was gegeben sein sollte, damit wir $\hat{\theta}$ als Schätzfunktion $\theta$ verwenden.

- Wir betrachten dafür folgende drei **Gütekriterien**:
    - Erwartungstreue
    - Effizienz
    - Konsistenz

# Erwartungstreue



- Zur Erinnerung: Je nachdem, welche konkrete Stichprobe zufällig gezogen wird, resultieren unterschiedliche Schätzwerte.
- Ein in einer konkreten Stichprobe berechneter Schätzwert wird daher in den seltensten Fällen genau dem wahren Parameterwert entsprechen.
- Wir können jedoch fordern, dass die Schätzwerte bei unendlicher Wiederholung des Zufallsexperiments – d.h. falls wir unendlich viele einfache Zufallsstichproben der Größe *n* ziehen würden – zumindest im Mittel dem wahren Parameterwert entsprechen sollten. Dieses Kriterium wird Erwartungstreue genannt.
- Da die Schätzwerte Realisationen der Schätzfunktion sind, und der Mittelwert unendlich vieler Realisationen einer Zufallsvariable deren Erwartungswert ist, ergibt sich:
- Eine Schätzfunktion $\hat{\theta}$ für einen Parameter $\theta$ ist **erwartungstreu**, falls gilt:

$$E(\hat{\theta}) = \theta$$

# Standardfehler einer Schätzfunktion I



- Eine zweite Anforderung an eine Schätzfunktion ist, dass sie zu möglichst „präzisen" oder „genauen" Schätzungen führen sollte.

- Wie bei der Erwartungstreue betrachten wir auch hier wieder den Fall, dass wir das zugrunde liegende Zufallsexperiment – also die Ziehung einer einfachen Zufallsstichprobe mit Umfang *n* – unendlich oft durchführen.

- Bei einer Schätzfunktion mit hoher „Genauigkeit" würden wir erwarten, dass ihre Realisationen – die Schätzwerte – bei wiederholter Ziehung der Stichprobe nur in geringerem Ausmaß um den wahren Parameterwert streuen.

- In diesem Fall liefern nämlich viele der möglichen zufälligen Stichproben einen Schätzwert, der nahe an dem wahren Parameterwert liegt.

# Standardfehler einer Schätzfunktion II



- Als Maß für die Genauigkeit einer Schätzfunktion kommen also prinzipiell alle Streuungsmaße für Zufallsvariablen infrage.
- Meistens wird die Standardabweichung $SD(\hat{\theta})$ der Schätzfunktion $\hat{\theta}$ betrachtet.
- Diese wird im Fall einer Schätzfunktion $\hat{\theta}$ **Standardfehler** $SE(\hat{\theta})$ von $\hat{\theta}$ genannt:

$$SE(\hat{\theta}) = SD(\hat{\theta})$$

- Je geringer der Standardfehler, desto weniger streuen die Schätzwerte und desto genauer ist die Schätzfunktion (Erwartungstreue vorausgesetzt).
- Bemerkung: SE kommt von **S**tandard **E**rror, der englischen Bezeichnung für den Standardfehler.

# Standardfehler einer Schätzfunktion III



- Der wahre Wert des Parameters sei $\theta = 10$

- Beispiel für die Schätzwerte, die aus einer erwartungstreuen Schätzfunktion $\hat{\theta}$ für $\theta$ mit geringem Standardfehler $SE(\hat{\theta})$ resultieren könnten:
  - Zufällige Stichprobe 1: $\hat{\theta}_{Wert} = 10.3$
  - Zufällige Stichprobe 2: $\hat{\theta}_{Wert} = 9.8$
  - Zufällige Stichprobe 3: $\hat{\theta}_{Wert} = 10.1$
  - usw.

  ➢ Geringe Streuung der Schätzwerte, die meisten Schätzwerte liegen nahe um den wahren Wert $\theta = 10$.

# Standardfehler einer Schätzfunktion IV



- Der wahre Wert des Parameters sei $\theta = 10$
- Beispiele für Schätzwerte, die aus einer erwartungstreuen Schätzfunktion $\hat{\theta}$ für $\theta$ mit hohem Standardfehler $SE(\hat{\theta})$ resultieren könnten:
  - Zufällige Stichprobe 1: $\hat{\theta}_{Wert} = 40$
  - Zufällige Stichprobe 2: $\hat{\theta}_{Wert} = -10$
  - Zufällige Stichprobe 3: $\hat{\theta}_{Wert} = -20$
  - usw.

  - ➢ Starke Streuung, die meisten Schätzwerte liegen weit weg vom wahren Wert.
  - ➢ Da die Schätzfunktion erwartungstreu ist, ist der Mittelwert der Schätzwerte unendlich vieler Stichproben immer noch gleich dem wahren Wert.
  - ➢ Dies hilft uns jedoch nicht wirklich, da wir in der Realität nur eine Stichprobe und daher nur einen Schätzwert $\hat{\theta}_{Wert}$ haben. Dieser kann sehr weit weg vom wahren Parameterwert liegen.

# Effizienz



- Nach diesen Vorüberlegungen können wir nun das Gütekriterium der Effizienz formulieren:

- Eine **erwartungstreue** Schätzfunktion ist **effizient**, falls sie einen geringeren Standardfehler als alle anderen erwartungstreuen Schätzfunktionen aufweist.

- Eine effiziente Schätzfunktion ist also die „präziseste" Schätzfunktion unter allen erwartungstreuen Schätzfunktionen.

- Die Effizienz zu beweisen ist sehr aufwändig. Wir gehen deshalb im Folgenden darauf immer nur sehr kurz ein.

# Konsistenz

- Eine weitere wünschenswerte Eigenschaft von Schätzfunktionen ist, dass sie mit wachsendem Stichprobenumfang *n* immer „präziser“ werden.

- Das heißt: mit wachsendem Stichprobenumfang *n* sollte der Standardfehler der Schätzfunktion immer kleiner werden und schließlich gegen Null gehen.

- Formal: Eine **erwartungstreue** Schätzfunktion $\hat{\theta}$ für einen Parameter $\theta$ ist **konsistent**, falls gilt:

$$\lim_{n \to \infty} SE(\hat{\theta}) = 0$$

# Gütekriterien unserer vorgeschlagenen Schätzfunktionen

# Gütekriterien für Schätzfunktionen



- Wir werden nun für die Schätzfunktionen aus unseren Beispielen jeweils
  - den Erwartungswert bestimmen
  - die Varianz und die Standardabweichung (d.h. den Standardfehler) bestimmen

- Damit können wir dann die einzelnen Schätzfunktionen anhand der Gütekriterien beurteilen.

# Schätzfunktion für den Parameter $\pi$ einer Bernoulli-Verteilung I



- Ausgangssituation: $X_1, X_2, \ldots, X_n$ mit $X_i \overset{\text{iid}}{\sim} Be(\pi)$

- Erste Idee für eine Schätzfunktion für $\pi$ :

$$\hat{\pi} = \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- Entsprechender Schätzwert für $\pi$ :

$$\hat{\pi}_{Wert} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

# Schätzfunktion für den Parameter $\pi$ einer Bernoulli-Verteilung II



- Erwartungswert von $\hat{\pi} = \bar{X}$:

Rechenregeln aus VL 5

$$E(\hat{\pi}) = E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \pi = \frac{1}{n}\cdot n \cdot \pi = \pi$$

Bekannt aus den Vorannahmen

- Varianz von $\hat{\pi} = \bar{X}$:

Rechenregeln aus VL 5

$$Var(\hat{\pi}) = Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \pi(1-\pi) = \frac{\pi(1-\pi)}{n}$$

Bekannt aus den Vorannahmen

- Standardfehler von $\hat{\pi} = \bar{X}$:

$$SE(\hat{\pi}) = SD(\hat{\pi}) = \sqrt{Var(\hat{\pi})} = \sqrt{\frac{\pi(1-\pi)}{n}}$$

# Schätzfunktion für den Parameter $\pi$ einer Bernoulli-Verteilung III



- Gütekriterien:

  - ✓ $\hat{\pi} = \bar{X}$ ist erwartungstreu (siehe vorherige Folie)
  - ✓ $\hat{\pi} = \bar{X}$ ist effizient (Beweis schwierig)
  - ✓ $\hat{\pi} = \bar{X}$ ist konsistent, da

$$\lim_{n \to \infty} SE(\hat{\pi}) = \lim_{n \to \infty} \sqrt{\frac{\pi(1-\pi)}{n}} = 0$$

➢ Da alle drei Gütekriterien erfüllt sind, können wir also $\bar{X}$ als Schätzfunktion für $\pi$ verwenden.

# Zusammenfassung Beispiel Depression I



- Wir interessieren uns für die relative Häufigkeit $h_D$ der Depression **in einer Population**.
- Wir ziehen eine einfache Zufallsstichprobe mit *n* = 100 Personen aus dieser Population.
- $X_1, X_2, \ldots, X_i, \ldots, X_{100}$ sind Bernoulli-Variablen, die jeweils den Wert 1 annehmen, falls die *i*-te Person in der Stichprobe depressiv ist.
- $X_i \overset{\text{iid}}{\sim} Be(\pi)$
- $\pi = h_D$

➢ Als Schätzfunktion für $\pi$ wählen wir $\hat{\pi} = \bar{X}$ , da wir wissen, dass diese Schätzfunktion eine erwartungstreue, effiziente und konsistente Schätzfunktion für $\pi$ ist.

# Zusammenfassung Beispiel Depression II



- Die Realisation der Schätzfunktion $\hat{\pi} = \bar{X}$ ist der konkrete Schätzwert $\hat{\pi}_{Wert} = \bar{x}$ .
  Er entspricht der relativen Häufigkeit der Depression **in unserer Stichprobe** und wir können ihn aus den einzelnen Realisationen $x_1, x_2, \ldots, x_{100}$ berechnen.

- Zum Beispiel könnte sich ergeben:

$$\hat{\pi}_{Wert} = \bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 0.2$$

- Unser Schätzwert für $\pi$ ist in diesem Fall $0.2$

- Da $\pi$ der relativen Häufigkeit $h_D$ in der Population entspricht, gehen wir auf Basis unserer Stichprobe davon aus, dass die relative Häufigkeit $h_D$ in der Population gleich 0.2 ist.

# Schätzfunktion für den Parameter $\mu$ einer Normal-Verteilung I



- Ausgangssituation: $X_1, X_2, \ldots, X_n$ mit $X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$

- Erste Idee für eine Schätzfunktion für $\mu$:

$$\hat{\mu} = \bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

- Entsprechender Schätzwert für $\mu$:

$$\hat{\mu}_{Wert} = \bar{x} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

# Schätzfunktion für den Parameter $\mu$ einer Normal-Verteilung II



- Erwartungswert von $\hat{\mu} = \bar{X}$:

Rechenregeln aus VL 5

$$E(\hat{\mu}) = E(\bar{X}) = E\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{1}{n}\sum_{i=1}^{n} \mu = \frac{1}{n} \cdot n \cdot \mu = \mu$$

Bekannt aus den Vorannahmen

- Varianz von $\hat{\mu} = \bar{X}$:

Rechenregeln aus VL 5

$$Var(\hat{\mu}) = Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2}\sum_{i=1}^{n} \sigma^2 = \frac{1}{n^2} \cdot n \cdot \sigma^2 = \frac{\sigma^2}{n}$$

Bekannt aus den Vorannahmen

- Standardfehler von $\hat{\mu} = \bar{X}$:

$$SE(\hat{\mu}) = SD(\hat{\mu}) = \sqrt{Var(\hat{\mu})} = \sqrt{\frac{\sigma^2}{n}}$$

# Schätzfunktion für den Parameter $\mu$ einer Normal-Verteilung III



- Gütekriterien:

  - ✓ $\hat{\mu} = \bar{X}$ ist erwartungstreu (siehe vorherige Folie)
  - ✓ $\hat{\mu} = \bar{X}$ ist effizient (Beweis schwierig)
  - ✓ $\hat{\mu} = \bar{X}$ ist konsistent, da

$$\lim_{n \to \infty} SE(\hat{\mu}) = \lim_{n \to \infty} \sqrt{\frac{\sigma^2}{n}} = 0$$

➢ Da alle drei Gütekriterien erfüllt sind, können wir also $\bar{X}$ als Schätzfunktion für $\mu$ verwenden.

# Ausblick: Wahrscheinlichkeitsverteilung von $\bar{X}$



- Ausgangssituation: $X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$ für *i* = 1, 2, …, *n* und $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

- Was ist dann die Wahrscheinlichkeitsverteilung $P_{\bar{X}}$ von $\bar{X}$ ?

- Man kann zeigen, dass unter der Voraussetzung, dass alle $X_i$ unabhängig und normalverteilt sind, auch $\bar{X}$ normalverteilt ist (Beweis sehr schwierig), d.h.:

$$\bar{X} \sim N(E(\bar{X}), Var(\bar{X}))$$

- Den Erwartungswert und die Varianz von $\bar{X}$ haben wir gerade bestimmt:

$$E(\bar{X}) = \mu$$

$$Var(\bar{X}) = \frac{\sigma^2}{n}$$

- Damit ergibt sich als Wahrscheinlichkeitsverteilung für $\bar{X}$:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

# Schätzfunktion für den Parameter $\sigma^2$ einer Normal-Verteilung I



- Ausgangssituation: $X_1, X_2, \ldots, X_n$ mit $X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$

- Erste Idee für eine Schätzfunktion für $\sigma^2$:

$$\hat{\sigma}^2 = S^2_{emp} = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- Entsprechender Schätzwert für $\sigma^2$:

$$\hat{\sigma}^2_{Wert} = s^2_{emp} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

# Schätzfunktion für den Parameter $\sigma^2$ einer Normal-Verteilung II



- Erwartungswert von $\hat{\sigma}^2 = S^2_{emp}$:

$$E(\hat{\sigma}^2) = E\left(S^2_{emp}\right) = \cdots = \frac{n-1}{n}\sigma^2 \neq \sigma^2$$

✗ Das heißt: $S^2_{emp}$ ist **nicht** erwartungstreu für $\sigma^2$!

➢ Wir hätten also lieber eine andere Schätzfunktion für $\sigma^2$.

# Schätzfunktion für den Parameter $\sigma^2$ einer Normal-Verteilung III



- Idee: Wegen

$$E\left(S_{emp}^2\right) = \frac{n-1}{n}\sigma^2$$

gilt

Rechenregeln aus VL 5

$$E\left(\frac{n}{n-1}S_{emp}^2\right) = \frac{n}{n-1}E\left(S_{emp}^2\right) = \frac{n}{n-1}\cdot\frac{n-1}{n}\sigma^2 = \sigma^2$$

- Also ist

$$\hat{\sigma}^2 = S^2 = \frac{n}{n-1}S_{emp}^2 = \frac{n}{n-1}\cdot\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

eine erwartungstreue Schätzfunktion für $\sigma^2$.

# Schätzfunktion für den Parameter $\sigma^2$ einer Normal-Verteilung IV



- Ausgangssituation: $X_1, X_2, \ldots, X_n$ mit $X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$

- Erwartungstreue Schätzfunktion für $\sigma^2$:

$$\hat{\sigma}^2 = S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$$

- Schätzwert für $\sigma^2$:

$$\hat{\sigma}^2_{Wert} = s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

# Schätzfunktion für den Parameter $\sigma^2$ einer Normal-Verteilung V



- Erwartungswert von $\hat{\sigma}^2 = S^2$:

$$E(\hat{\sigma}^2) = E(S^2) = \sigma^2$$

- Varianz von $\hat{\sigma}^2 = S^2$:

$$Var(\hat{\sigma}^2) = Var(S^2) = \cdots = \frac{2\sigma^4}{n-1}$$

- Standardfehler von $\hat{\sigma}^2 = S^2$:

$$SE(\hat{\sigma}^2) = SD(\hat{\sigma}^2) = \sqrt{Var(\hat{\sigma}^2)} = \sqrt{\frac{2\sigma^4}{n-1}}$$

# Schätzfunktion für den Parameter $\sigma^2$ einer Normal-Verteilung VI



- Gütekriterien:

  - ✓ $\hat{\sigma}^2 = S^2$ ist erwartungstreu (siehe vorherige Folie)
  - ✓ $\hat{\sigma}^2 = S^2$ ist effizient (Beweis schwierig)
  - ✓ $\hat{\sigma}^2 = S^2$ ist konsistent, da

$$\lim_{n\to\infty} SE(\hat{\sigma}^2) = \lim_{n\to\infty} \sqrt{\frac{2\sigma^4}{n-1}} = 0$$

➢ Im Gegensatz zu $S^2_{emp}$ (nicht erwartungstreu) erfüllt $S^2$ alle Gütekriterien. Wir verwenden deshalb $S^2$ als Schätzfunktion für $\sigma^2$.

# Schätzfunktion für den Parameter $\sigma^2$ einer Normal-Verteilung VII



- Bemerkung: Die Erwartungstreue von $S^2$ für $\sigma^2$ ist der Grund dafür, dass die Funktion `var()` in R die Varianz einer Variablen nach der Formel

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

  und nicht nach der Formel

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

  berechnet.

- Das Gleiche gilt auch für die R-Funktionen `sd()`, und `cov()`. Auch in diesen Fällen teilt R immer durch $n-1$, da die entsprechende Funktion nur dann erwartungstreu für die Schätzung des entsprechenden Parameters in einer Population ist.
- Der Unterschied zwischen den erwartungstreuen und nicht erwartungstreuen Versionen ist bei Stichproben mit angemessener Größe vernachlässigbar. Daher ist es in der Praxis üblich, die entsprechenden R Funktionen auch dann zu benutzen, wenn man sich nur für die entsprechenden deskriptivstatistischen Maße interessiert. Trotzdem ist es wichtig, konzeptuell den Unterschied zu verstehen.

# Zusammenfassung Beispiel Intelligenz I



- Wir interessieren uns für den Mittelwert $\bar{x}_{IQ}$ und die empirische Varianz $s^2_{emp\ IQ}$ des IQ **in einer Population**.
- Wir gehen davon aus, dass das Histogramm des IQ in der Population durch die Wahrscheinlichkeitsdichte einer Normalverteilung approximiert werden kann.
- Wir ziehen eine einfache Zufallsstichprobe mit *n* = 100 Personen aus dieser Population.
- $X_1, X_2, \ldots, X_i, \ldots, X_{100}$ sind Zufallsvariablen, die jeweils für den IQ-Wert der *i*-ten Person in der Stichprobe stehen.
- $X_i \overset{\text{iid}}{\sim} N(\mu, \sigma^2)$
- $\mu = \bar{x}_{IQ}$ und $\sigma^2 = s^2_{emp\ IQ}$
- Als Schätzfunktionen für $\mu$ und $\sigma^2$ wählen wir $\hat{\mu} = \bar{X}$ und $\hat{\sigma}^2 = S^2$, da wir wissen, dass diese Schätzfunktionen jeweils erwartungstreue, effiziente und konsistente Schätzfunktionen für $\mu$ und $\sigma^2$ sind.

# Zusammenfassung Beispiel Intelligenz II



- Die Realisationen der Schätzfunktionen $\hat{\mu} = \bar{X}$ und $\hat{\sigma}^2 = S^2$ sind die konkreten Schätzwerte $\hat{\mu}_{Wert} = \bar{x}$ und $\hat{\sigma}^2_{Wert} = s^2$. Diese können wir aus den einzelnen Realisationen $x_1$, $x_2$, …, $x_n$ berechnen.

- Zum Beispiel könnte sich ergeben:

$$\hat{\mu}_{Wert} = \bar{x} = 102$$
$$\hat{\sigma}^2_{Wert} = s^2 = 220$$

- Unser Schätzwert für $\mu$ ist in diesem Fall 102 und unser Schätzwert für $\sigma^2$ ist 220.

- Da $\mu$ dem Mittelwert $\bar{x}_{IQ}$ in der Population entspricht und $\sigma^2$ der empirischen Varianz $s^2_{emp\ IQ}$ in der Population, gehen wir auf Basis unserer Stichprobe davon aus, dass $\bar{x}_{IQ}$ gleich 102 ist und $s^2_{emp\ IQ}$ gleich 220.

# Überblick
# Zufallsstichprobe und Punktschätzung



- **Einfache Zufallsstichprobe**
  - *n* Personen werden **unabhängig** voneinander gezogen
  - Jede Person hat die gleiche Wahrscheinlichkeit gezogen zu werden
  - Grundlage für iid Zufallsvariablen

- **Schätzfunktionen** sind Zufallsvariablen,
  - für die wir Erwartungswert und Standardabweichung (Standardfehler) berechnen können
  - die idealerweise erwartungstreu, effizient und konsistent sein sollten
  - mit deren Hilfe wir Parameter (gesuchter Parameter einer Verteilung entspricht interessierender Größe in der Population) aus Stichprobendaten schätzen