

## Statistik II

Prof. Dr. Simone Abendschön  
Vorlesung am 8.5.23

- **MAP-Anmeldung in flexnow geöffnet**
  - Normalfall: Sie müssen sich extra einflexen um teilzunehmen!
  - Ausnahme alte PO: Sie sind automatisch zur Klausur angemeldet

- **Achtung: umfassen teilweise Stoff von 2 Vorlesungen (Hintergrund: andere Vorlesungsstruktur in Coronasemester)**
- **Abgleich mit VI-Folien**
- **Insgesamt um die 6-7 Videos**

# Plan heute

- Exkurs Integrationsbarometer
- Hintergrund Inferenzstatistik (Warum müssen wir uns damit beschäftigen?)
- Zufallsstichprobe und Wahrscheinlichkeit
- Kurzer Exkurs Integrationsbarometer
- Wahrscheinlichkeiten für diskrete und kontinuierliche Zufallsvariablen
- Rolle der (Standard-) Normalverteilung
- Übungsbeispiele

# Lernziele

- Sie erweitern Ihre Kenntnisse über die sog. „Normalverteilung“ und wissen wozu sie in der Statistik dient
- Sie können Flächenanteile und damit Wahrscheinlichkeiten innerhalb der Normalverteilung berechnen

# Beispiel Integrationsbarometer des SVR

Tab. 1 Eckdaten zum SVR-Integrationsbarometer 2018

Grundgesamtheit	Bevölkerung ohne und mit Migrationshintergrund in Deutschland ab 15 Jahren
Art der Befragung	telefonische Interviews (CATI)
realisierte Stichprobe	9.298 Fälle
Erhebungszeitraum	19.07.2017 – 31.01.2018
Stichprobendesign	Dual-Frame; disproportionale Schichtung nach Herkunftsgruppen und Bundesländern
Auswahlgrundlagen	ADM-Telefonauswahlgrundlage 2017 Festnetz und Mobilnetz mit den Schichten Standard- und Auslandstarife, zusätzlich für spezielle Sprachgruppen onomastisch markierte Telefonbucheinträge

# Beispiel Integrationsbarometer des SVR

Sachverständigenrat deutscher Stiftungen  
für Integration und Migration



SACHVERSTÄNDIGENRAT

FORSCHUNGSBEREICH

JAHRESGUTACHTEN

BAROMETER

PUBLIKATIONEN

POSITIONEN

THEMEN KURZ & BÜNDIG

VERANSTALTUNGEN

Die Mitglieder des  
Sachverständigenrats

Weitere Informationen >>



- Aufgaben und Ziele SVR
  - Unabhängige Politikberatung u.a. durch Forschungsprojekte rund um das Thema Migration und Integration
- Integrationsbarometer:
  - Alle 2-3 Jahre
  - repräsentative Bevölkerungsumfrage von in Deutschland lebenden Menschen mit und ohne Migrationshintergrund
  - „Integrationsklima“ messen
  - „beide Seiten der Einwanderungsgesellschaft“



- Unverzichtbarer Bestandteil jeder empirischen Studie
  - Kriterien empirischer Sozialforschung: Offenheit, Transparenz, Nachvollziehbarkeit, prinzipielle Unabgeschlossenheit von Forschung
- Angaben zur Stichprobenziehung, Datenerhebung, etc.
- Grundlage von Bewertung und Weiterentwicklung: Gütekriterien Objektivität, Reliabilität
- Empirische Abschlussarbeit: Kapitel Methoden

# Beispiel Integrationsbarometer des SVR

Tab. 1 Eckdaten zum SVR-Integrationsbarometer 2018

Grundgesamtheit	Bevölkerung ohne und mit Migrationshintergrund in Deutschland ab 15 Jahren
Art der Befragung	telefonische Interviews (CATI)
realisierte Stichprobe	9.298 Fälle
Erhebungszeitraum	19.07.2017 – 31.01.2018
Stichprobendesign	Dual-Frame; disproportionale Schichtung nach Herkunftsgruppen und Bundesländern
Auswahlgrundlagen	ADM-Telefonauswahlgrundlage 2017 Festnetz und Mobilnetz mit den Schichten Standard- und Auslandstarife, zusätzlich für spezielle Sprachgruppen onomastisch markierte Telefonbucheinträge

## Beispiel Integrationsbarometer des SWR

- CATI
- ADM
- Dual Frame
- Disproportionale geschichtete Zufallsauswahl
- Onomastische Markierung Telefonbucheinträge

- Papierfragebogen (Paper And Pencil Interview, PAPI):
- Computerunterstützt:
  - Computer Assisted Personal Interview (CAPI)
  - **Computer Assisted Telephone Interview (CATI)**
  - Computer Assisted Self Interview (CASI)
- Online-Survey

- Telefonlabor mit technischer Ausstattung und entsprechend vielen Sitzplätzen (CATI-Labor)
  - Umfrageinstitute
  - PCs
  - Headsets mit hochempfindlichen Mikrofonen
  - Schnelle digitale Verbindungen
- Grundlegende Anforderungen an Interviewende:  
Gute Telefonstimme, Motivation
  - Technische und kommunikative Schulungen vor erstem Interview und Aufbauschulungen
  - Supervision



# Disproportionale Stichprobenziehung

- Gewährleistet, dass bestimmte „Untergruppen“ der Grundgesamtheit (GG) mit einer ausreichend großen Fallzahl in der Stichprobe sind
  - Allbus: „neue“ Bundesländer
  - Integrationsbarometer:  
Migrationshintergrund

- Arbeitskreis Deutscher Markt- und Sozialforschungsinstitute e.V.
- Stellen Stichprobenpläne für bevölkerungsrepräsentative Befragungen bereit (Face-to-Face-Interviews und Telefon)
- ADM-Design als mehrstufige Zufallsauswahl



## Dual Frame

- Für Telefonstichproben relevant
- Sowohl Festnetz- als auch Mobilfunknummern sind in der Auswahlgesamtheit
- Festnetz: aus registrierten Telefonnummern, teilweise zufällig erzeugten Nummern
- alle in der BRD nutzbaren Mobilfunknummern auf Basis von 10.000er Blöcken generiert werden (sogenannte Mobilfunkdatei); Inlands- und Auslandstarife



- Onomastik: Namensforschung
  - Onomastisches Quellenverzeichnis ordnet Herkunft auf Basis von Namen zu (für deutsche, japanische und türkische Herkunft sehr gute Trefferquote)
  - Integrationsbarometer: Migrationshintergrund

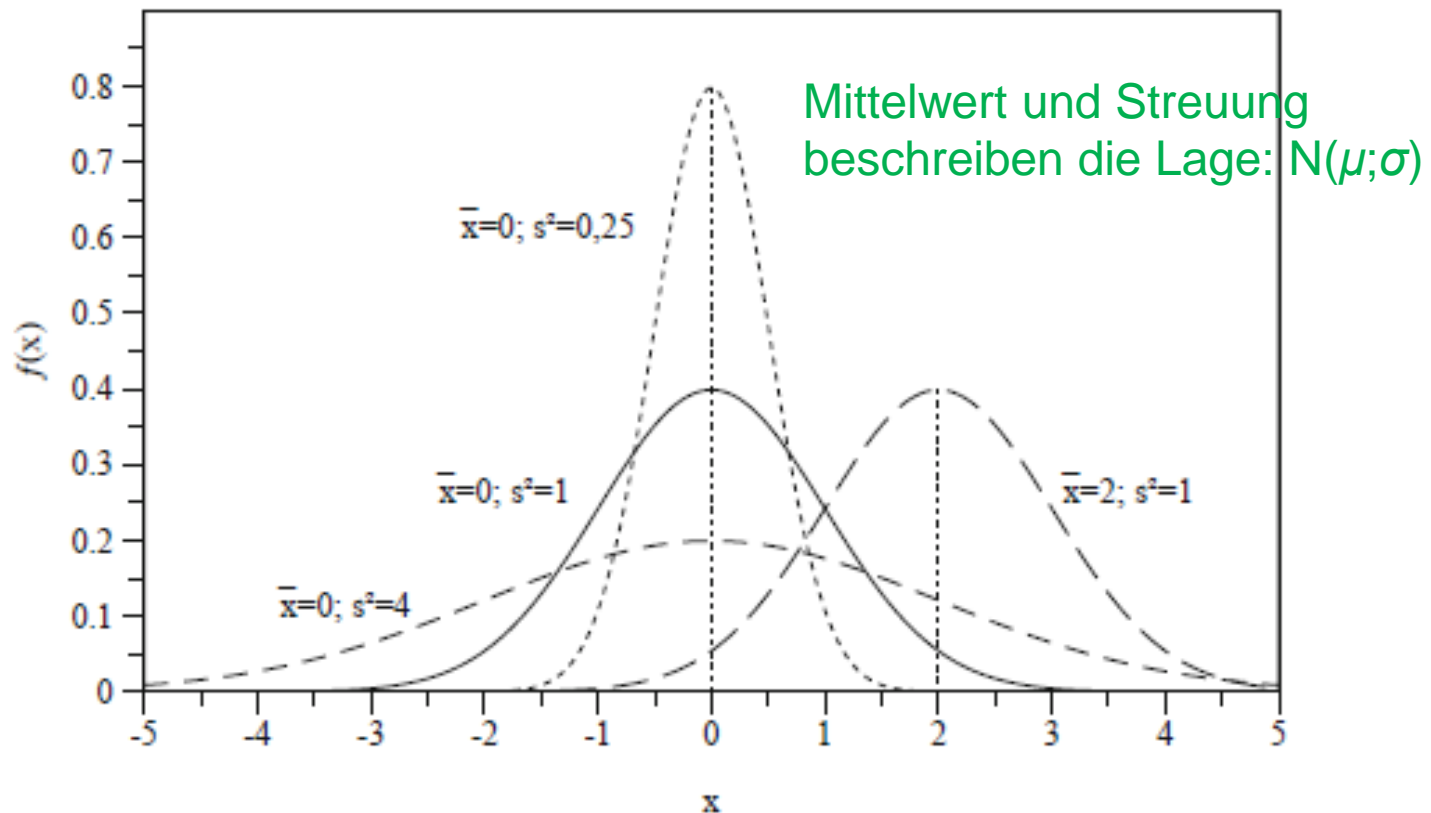
# Plan heute

- Exkurs Integrationsbarometer
- Hintergrund Inferenzstatistik (Warum müssen wir uns damit beschäftigen?)
- Zufallsstichprobe und Wahrscheinlichkeit
- Kurzer Exkurs Integrationsbarometer
- Wahrscheinlichkeiten für diskrete und kontinuierliche Zufallsvariablen
- Rolle der (Standard-) Normalverteilung
- Übungsbeispiele

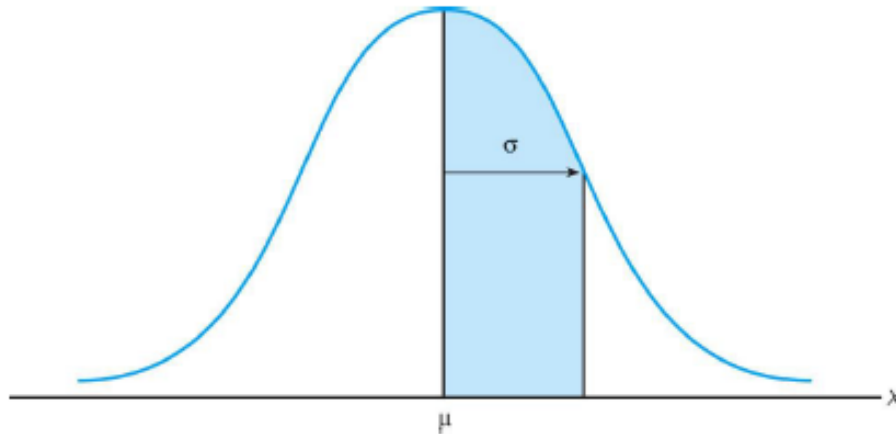
# Wdh. Normalverteilung

- Deskriptiv: symmetrische Verteilungsform („Glocke“): Werte konzentrieren sich in der Mitte, treten mit größerem Abstand zur Mitte immer seltener auf
- Im „wirklichen Leben“: einige Merkmale treten normalverteilt in der Bevölkerung auf (IQ, Körpergröße)
- Inferenzstatistik:
  - zentrales Modell für Wahrscheinlichkeitsverteilungen für kontinuierliche Zufallsvariablen, sog. „stetige Verteilungen“
  - Stichprobenkennwerte sind (unter bestimmten Bedingungen) normalverteilt

Abbildung 10.4: Normalverteilungen mit verschiedenen Parametern  $\bar{x}$  und  $s^2$

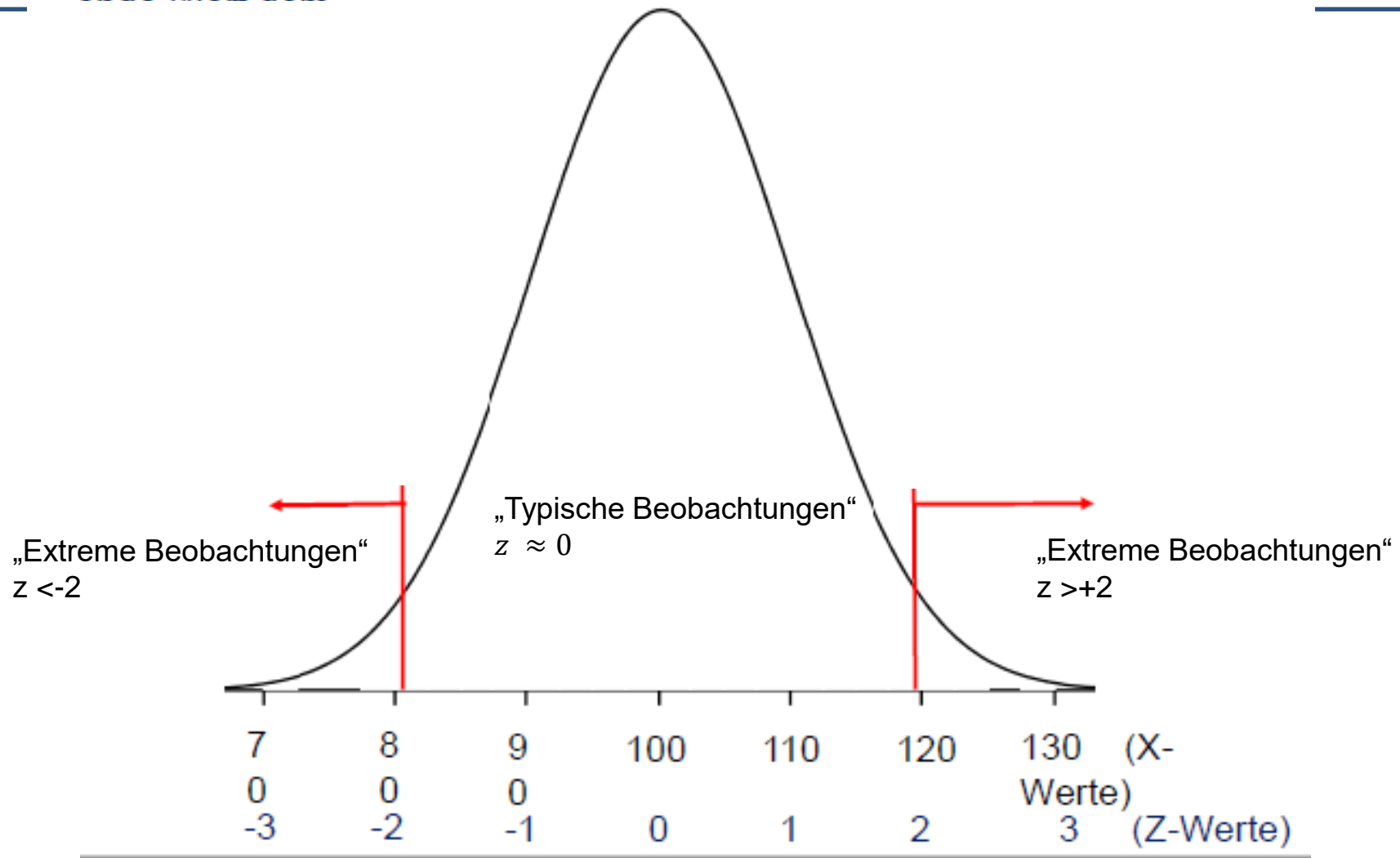


# Normalverteilung



- Symmetrisch
- Mittelwert=Median=Modus
- Größte Häufigkeiten in der Mitte, geringere Häufigkeiten rechts/links von der Mitte

# Z-Transformation → Standardnormalverteilung

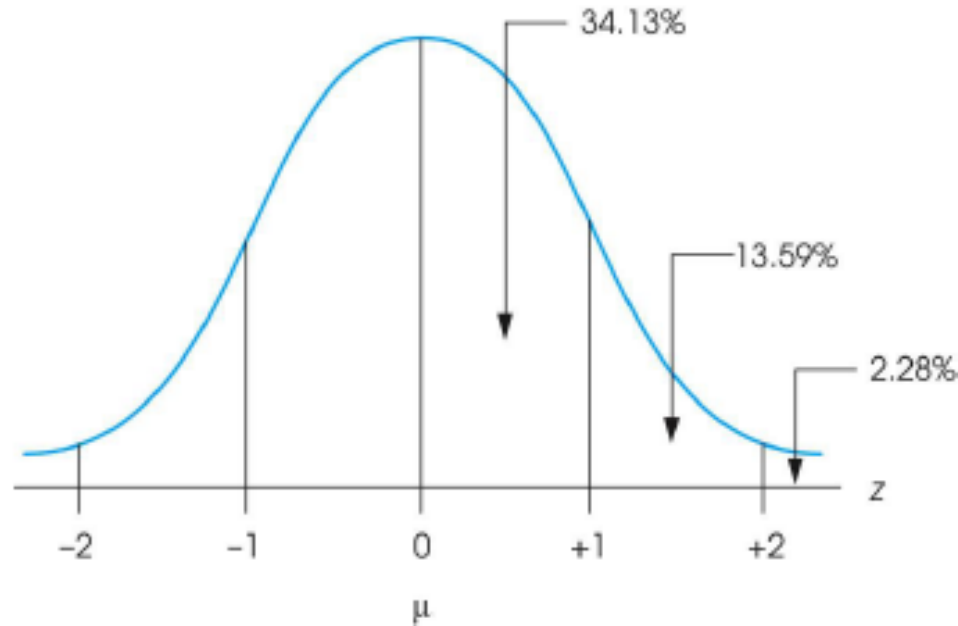


- **Standardnormalverteilung** (z-Transformation):  $\mu = 0$  und  $\sigma = 1$

# Normalverteilungsmodell

- Wahrscheinlichkeiten für kontinuierliche Variablen können nicht direkt berechnet werden
- Stattdessen: Wie wahrscheinlich ist es, dass eine Zufallsvariable in ein bestimmtes Intervall fällt
- Fläche unter der Kurve ist 1 (bzw. 100%)
- Rechnerische Bestimmung einzelner Werte/Intervalle sehr aufwendig → z-Tabelle, Statistikprogramme werden genutzt

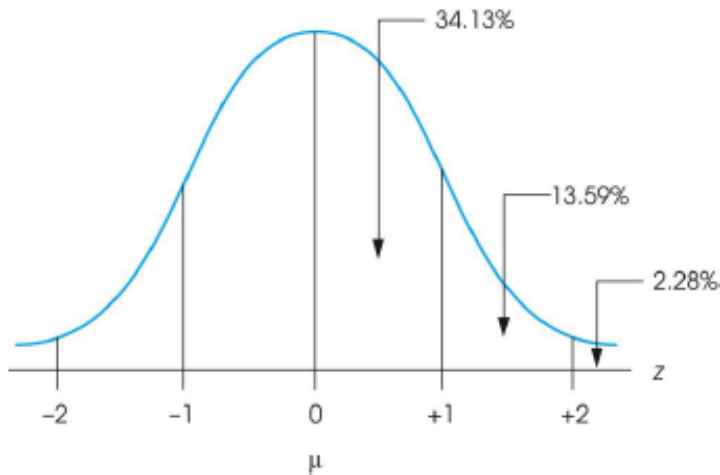
# Flächenanteile & Wahrscheinlichkeiten für z-Werte



- Wahrscheinlichkeitsdichte für Werte zwischen  $-\infty$  und  $+\infty$  Fläche unter der Kurve = 1, d.h. 100%
- Wahrscheinlichkeit für Wert aus einem bestimmten Bereich = Fläche über diesem Intervall → Bestimmte Intervalle entsprechen bestimmten Flächenanteilen



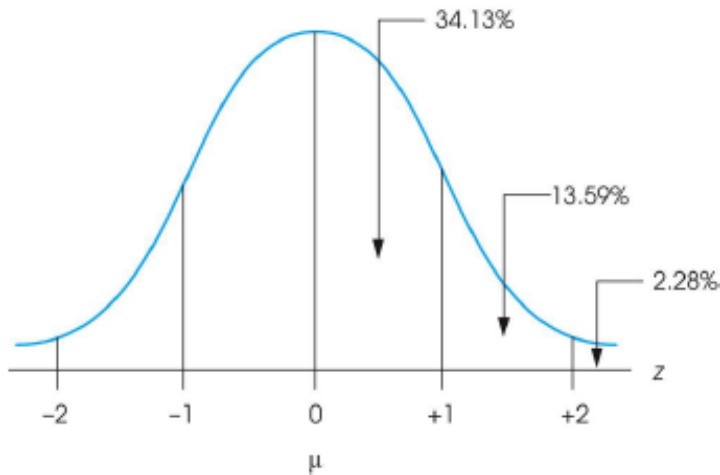
# Flächenanteile & Wahrscheinlichkeiten für z-Werte



Intervall	Flächenanteil
$[\mu - 1 \cdot \sigma; \mu + 1 \cdot \sigma]$	68.3%
$[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$	95.4%
$[\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma]$	99.7%

- Häufigkeiten/Wahrscheinlichkeiten werden durch Flächen repräsentiert
- 100% aller Fälle liegen unter der Normalverteilungskurve
- Bestimmte Intervalle entsprechen bestimmten Flächenanteilen

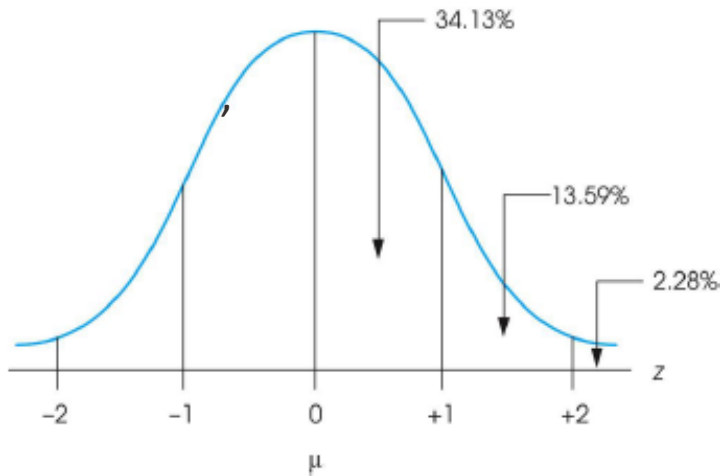
# Flächenanteile & Wahrscheinlichkeiten für z-Werte



Intervall	Flächenanteil
$[\mu - 1 \cdot \sigma; \mu + 1 \cdot \sigma]$	68.3%
$[\mu - 1.96 \cdot \sigma; \mu + 1.96 \cdot \sigma]$	95%
$[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$	95.4%
$[\mu - 2.58 \cdot \sigma; \mu + 2.58 \cdot \sigma]$	99.0%
$[\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma]$	99.7%

- Häufigkeiten werden durch Flächen repräsentiert
- 100% aller Fälle liegen unter der Normalverteilungskurve
- Bestimmte Intervalle entsprechen bestimmten Flächenanteilen

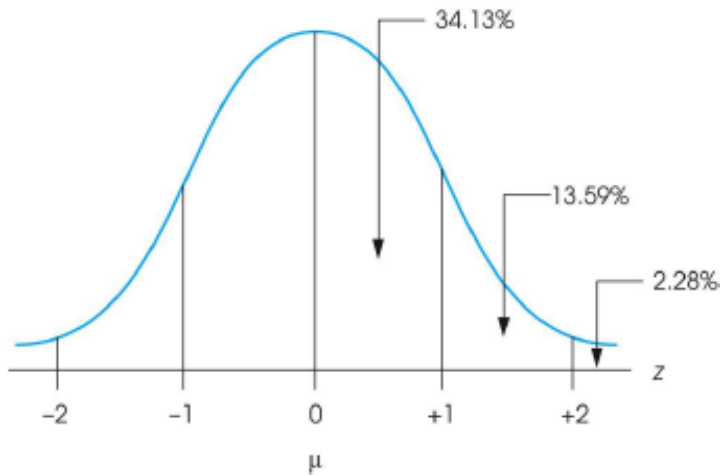
# Beispiel 1



Intervall	Flächenanteil
$[\mu - 1 \cdot \sigma; \mu + 1 \cdot \sigma]$	68.3%
$[\mu - 1.96 \cdot \sigma; \mu + 1.96 \cdot \sigma]$	95%
$[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$	95.4%
$[\mu - 2.58 \cdot \sigma; \mu + 2.58 \cdot \sigma]$	99.0%
$[\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma]$	99.7%

Gegeben sei für eine Population von  $N = 50.000$  Personen deren Körpergröße (in cm) mit  $N(175;5)$  – wie groß sind 95% aller Personen?

# Flächenanteile & Wahrscheinlichkeiten für z-Werte



Intervall	Flächenanteil
$[\mu - 1 \cdot \sigma; \mu + 1 \cdot \sigma]$	68.3%
$[\mu - 1.96 \cdot \sigma; \mu + 1.96 \cdot \sigma]$	95%
$[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$	95.4%
$[\mu - 2.58 \cdot \sigma; \mu + 2.58 \cdot \sigma]$	99.0%
$[\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma]$	99.7%

- Wahrscheinlichkeiten/Häufigkeiten werden durch Flächen repräsentiert
- 100% aller Fälle liegen unter der Normalverteilungskurve
- Bestimmte Intervalle entsprechen bestimmten Flächenanteilen

## Beispiel 1

Intervall	Flächenanteil
$[\mu - 1 \cdot \sigma; \mu + 1 \cdot \sigma]$	68.3%
$[\mu - 1.96 \cdot \sigma; \mu + 1.96 \cdot \sigma]$	95%
$[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$	95.4%
$[\mu - 2.58 \cdot \sigma; \mu + 2.58 \cdot \sigma]$	99.0%
$[\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma]$	99.7%

Gegeben sei für eine Population von  $N = 50.000$  Personen deren Körpergröße (in cm) mit  $N(175;5)$  – wie groß sind 95% aller Personen?

# Beispiel 1

Intervall	Flächenanteil
$[\mu - 1 \cdot \sigma; \mu + 1 \cdot \sigma]$	68.3%
$[\mu - 1.96 \cdot \sigma; \mu + 1.96 \cdot \sigma]$	95%
$[\mu - 2 \cdot \sigma; \mu + 2 \cdot \sigma]$	95.4%
$[\mu - 2.58 \cdot \sigma; \mu + 2.58 \cdot \sigma]$	99.0%
$[\mu - 3 \cdot \sigma; \mu + 3 \cdot \sigma]$	99.7%

Gegeben sei für eine Population von  $N = 50.000$  Personen deren Körpergröße (in cm) mit  $N(175;5)$  – wie groß sind 95% aller Personen?

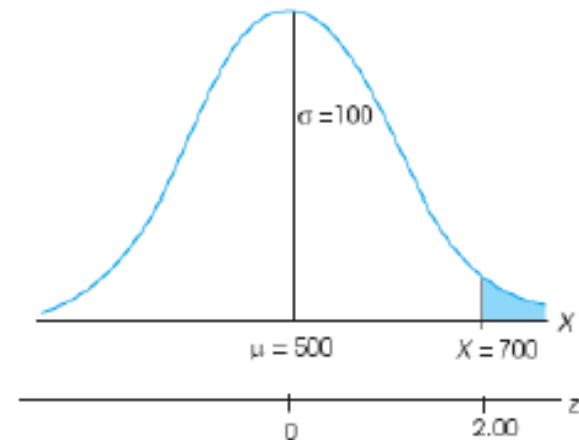
$$[175 - 1,96 \cdot 5; 175 + 1,96 \cdot 5] = [165,2; 184,8]$$

**„95% aller Personen haben eine Körpergröße zwischen 165,2cm und 184,8cm“**

## Beispiel 2

- Gegeben sei eine Verteilung in der Population mit  $\mu = 500$  und  $\sigma = 100$
- Wie hoch ist die Wahrscheinlichkeit, aus dieser Population zufällig ein Individuum mit einem höheren Wert als 700 zu ziehen („sampeln“)?

$$P(X > 700) = ?$$



## Beispiel 2

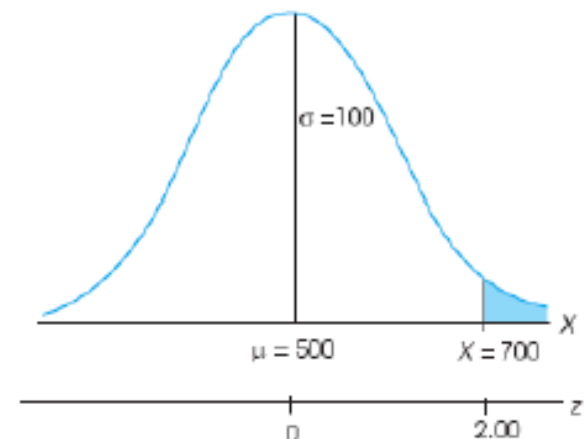
- Gegeben sei eine Verteilung in der Population mit  $\mu = 500$  und  $\sigma = 100$
- Wie hoch ist die Wahrscheinlichkeit, aus dieser Population zufällig einen Fall mit einem höheren Wert als 700 zu ziehen („sampeln“)?

$$P(X > 700) = ?$$

1. Welcher (Flächen-)Anteil ist größer als 700?
2. Exakte Position von X durch z-Wert bestimmen:

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

$$3. P(z > 2) =$$





## Beispiel 2

Gegeben sei eine Verteilung in der Population mit  $\mu = 500$  und  $\sigma = 100$ .

Wie hoch ist die Wahrscheinlichkeit, aus dieser Population zufällig ein Individuum mit einem höheren Wert als 700 zu ziehen/auszuwählen („sampeln“)?

$$P(X > 700) = ?$$

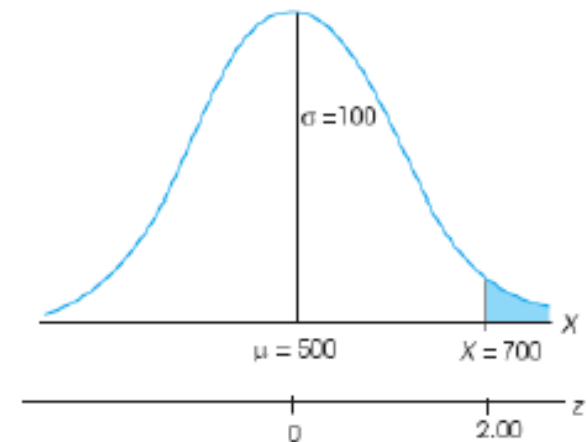
1. Welcher (Flächen-)Anteil ist größer als 700?

2. Exakte Position von X durch z-Wert bestimmen:

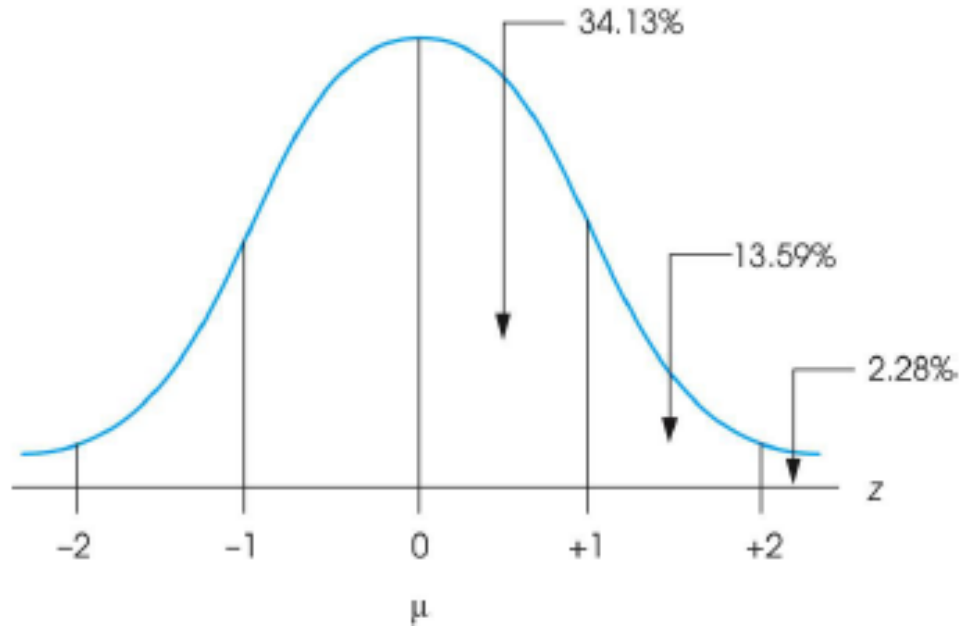
$$z = (700 - 500) / 100 = 2.0$$

$$3. P(z > 2) =$$

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



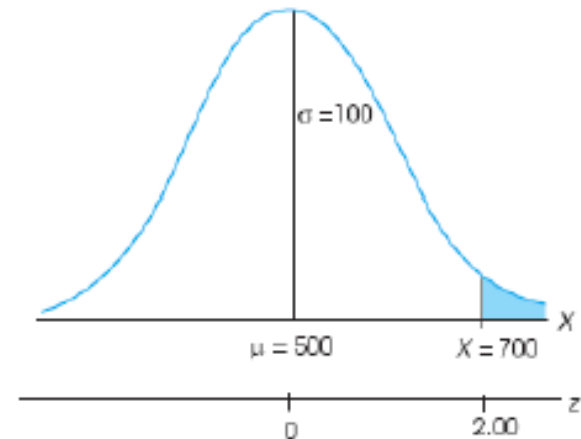
# Flächenanteile/Standardnormalverteilung



- Wahrscheinlichkeiten/ Häufigkeiten werden durch Flächen repräsentiert
- 100% aller Fälle liegen unter der Normalverteilungskurve
- Bestimmte Intervalle entsprechen bestimmten Flächenanteilen

# Wahrscheinlichkeit und Normalverteilung

Beispiel 2: Wie hoch ist die Wahrscheinlichkeit, jdn zufällig mit einem Wert größer 700 zu Ziehen?

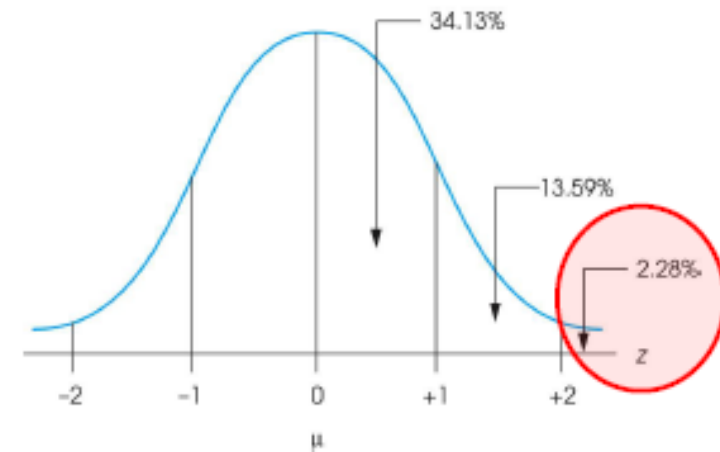


$$P(X > 700) = ?$$

1. Welcher (Flächen-)Anteil ist größer als 700?
2. Exakte Position von  $X$  durch  $z$ -Wert bestimmen

$$z = (700 - 500) / 100 = 2.0$$

$$3. P(z > 2) = 2,28\%$$



## Flächenanteile & Wahrscheinlichkeiten für z-Werte

- Berechnung der Flächenanteile und damit der Wahrscheinlichkeiten für kontinuierliche Variablen sehr aufwendig → für viele Verteilungen (auch Standardnormalverteilung) entsprechende Tabellen
- Statistikprogramme berechnen die Flächenanteile
- Z-Tabelle: Typischerweise sind die Flächen **links** vom Wert der Variablen tabelliert.

# z-Tabelle: siehe Formelsammlung

## 11 z-Tabelle

Die Tabelle enthält z-Werte, die auf zwei Stellen hinter dem Komma gerundet sind: z.B. -2.03, 1.07 oder 1.96.

**Leseübung:** Die Tabelle ist in zwei Teile aufgeteilt und somit ist auch jeder z-Wert in zwei Teile aufgeteilt: *Teil 1* mit der ersten Nachkommastelle (Spalte 1) und *Teil 2* mit der zweiten Nachkommastelle (alle folgenden Spalten). Jetzt suchen wir die Wahrscheinlichkeit (als Wert der Funktion  $\Phi_{0,1}(z)$ ), dass maximal ein z-Wert von -1.44 auftritt: In *Teil 1* geht man zur Zeile -1.4 und in dieser Zeile dann in die Spalte mit der zweiten Nachkommastelle 0.04 (*Teil 2*). Die gesuchte Wahrscheinlichkeit beträgt  $\Phi_{0,1}(z) = 0.0749$ . Das heißt, die Wahrscheinlichkeit das ein z-Wert kleiner gleich -1.44 ist beträgt 7.49 % (bzw. grafisch: die Fläche bis zum z-Wert von -1.44 beträgt 7.49 % der gesamten Fläche).

z-Wert	-0	-1	-2	-3	-4	-5	-6	-7	-8	-9
-2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
-2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
-2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
-2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
-2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
-2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
-2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
-2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
-2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
-2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
-1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
-1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294

# z-Tabelle

z-Wert	--0	--1	--2	--3	--4	--5	--6	--7	--8	--9
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	0.7703	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

Um Wert RECHTS von z=2 zu  
berechnen:  $1 - 0,9772 = 0,0228 \rightarrow 2,28\%$

## Beispiel 2

- Gegeben sei eine Verteilung in der Population mit  $\mu = 500$  und  $\sigma = 100$
- Wie hoch ist die Wahrscheinlichkeit, aus dieser Population zufällig ein Individuum mit einem höheren Wert als 700 zu ziehen („sampeln“)?

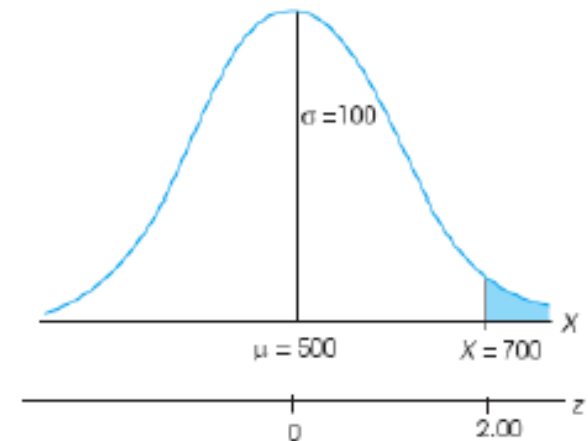
$$P(X > 700) = ?$$

1. Welcher (Flächen-)Anteil ist größer als 700?
2. Exakte Position von X durch z-Wert bestimmen:

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

$$3. P(z > 2) =$$

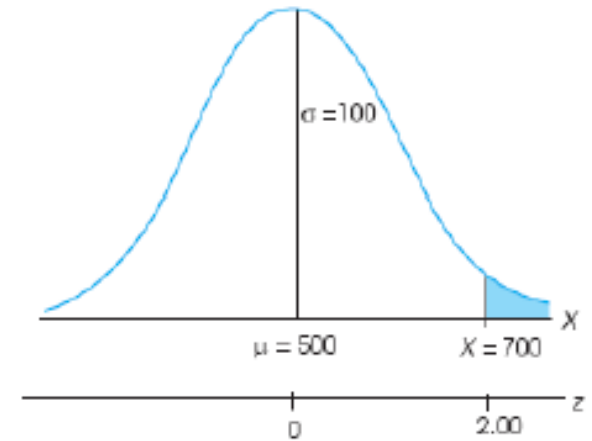
-2,0.	0,0228	0,022
-1,9.	0,0287	0,028
-1,8.	0,0359	0,035



# Wahrscheinlichkeit und Normalverteilung

Beispiel Wh.:

Gegeben sei eine Verteilung in der Population mit  $\mu = 500$  und  $\sigma = 100$ . Wie hoch ist die Wahrscheinlichkeit, aus dieser Population zufällig ein Individuum mit einem höheren Wert als 700 zu ziehen/auszuwählen („sampeln“)

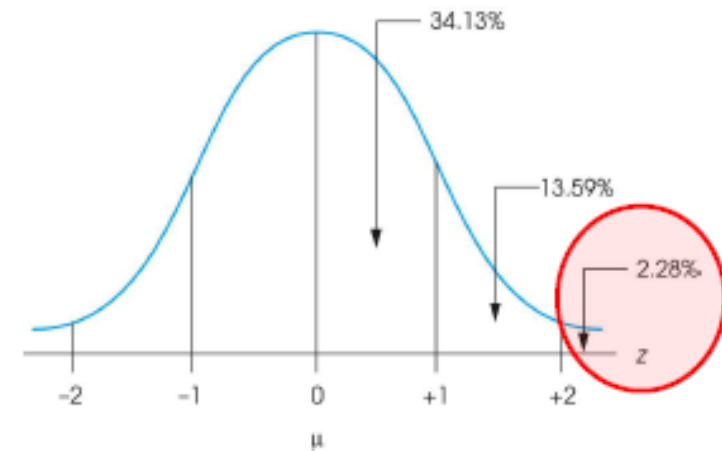


$$P(X > 700) = ?$$

1. Welcher (Flächen-)Anteil ist größer als 700?
2. Exakte Position von X durch z-Wert bestimmen:

$$z = (700 - 500) / 100 = 2.0$$

$$3. P(z > 2) = 2,28\%$$

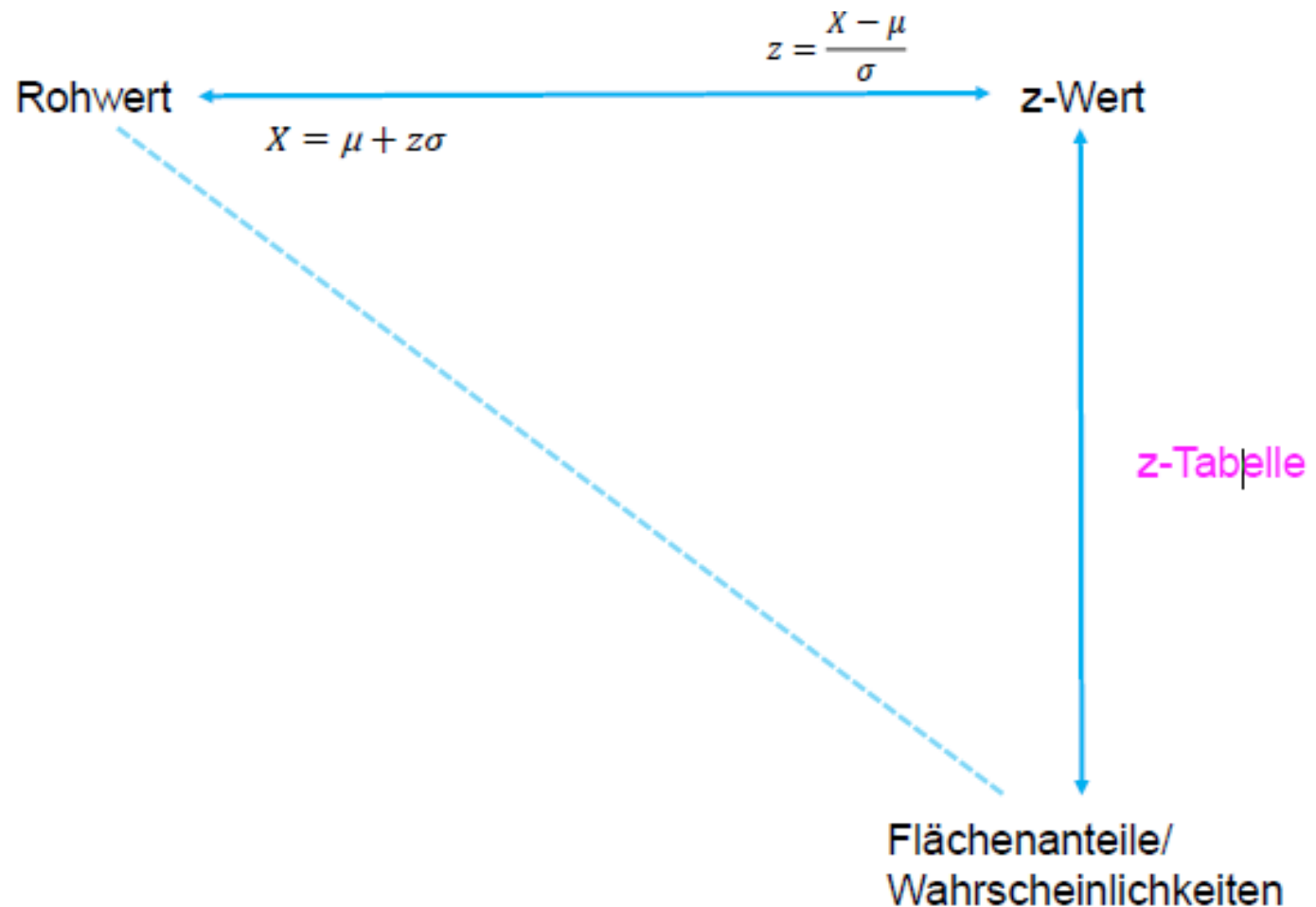




# Flächenanteile und Wahrscheinlichkeiten für z-Werte

- z-Werte-Tabelle (z-Tabelle, unitnormal table) enthält Anteile für alle z-Werte; Typischerweise sind die Flächen **links** vom jeweiligen z-Wert tabelliert.
- Anhand der Flächenanteile können die z-Werte bestimmt werden
- Wahrscheinlichkeit äquivalent zu den Flächenanteilen

# Flächenanteile & Wahrscheinlichkeiten für z-Werte



## Übungsbeispiel 1)

Welcher Flächenanteil der Normalverteilung entspricht z-Werten  $>1$ ?

Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen z-Wert  $> 1.0$  zu erhalten?

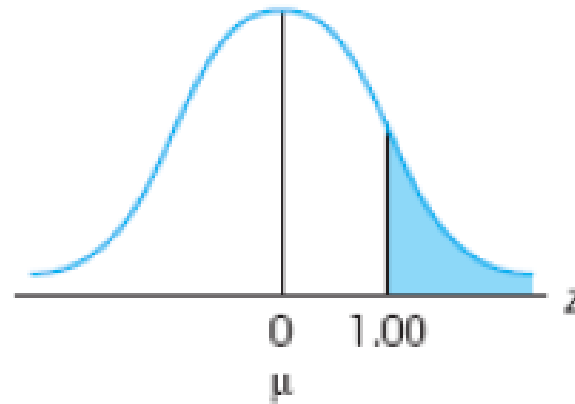
# Flächenanteile & Wahrscheinlichkeiten für z-Werte

z-Werte für Flächenanteile/Wahrscheinlichkeiten  
bestimmen:

Praktische Vorgehensweise:

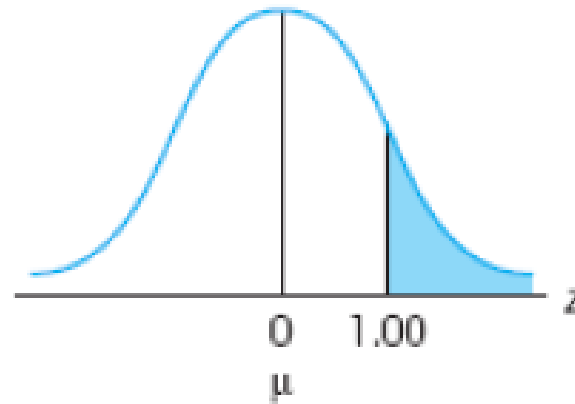
- Zunächst Normalverteilung mit der gesuchten Fläche skizzieren
- Dann entsprechende Werte aus z-Tabelle auswählen

## Übungsbeispiel 1)



Welcher Flächenanteil der Normalverteilung entspricht  $z$ -Werten  $>1$ ?  
Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen  $z$ -Wert  $> 1.0$  zu erhalten?

## Übungsbeispiel 1)



Welcher Flächenanteil der Normalverteilung entspricht z-Werten  $>1$ ?  
Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen z-Wert  $> 1.0$  zu erhalten?

Vorgehen:

- Skizzieren NV und gesuchte Fläche
- Bestimme  $z = 1.00$  in der z-Werte Tabelle: 0,8413
- **$P(z > 1.0) = 1 - 0,8413 = 0.1587 = 15.87\%$**

## Übungsbeispiel 2)

Welcher Flächenanteil der Normalverteilung entspricht z-Werten  $< 1,5$ ?

Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen z-Wert  $< 1.5$  zu erhalten?

# Flächenanteile & Wahrscheinlichkeiten für z-Werte

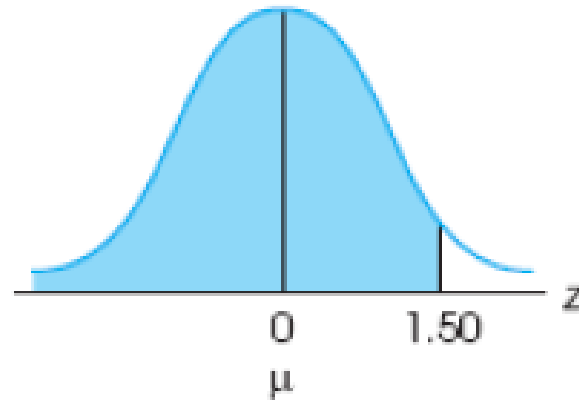
z-Werte für Flächenanteile/Wahrscheinlichkeiten  
bestimmen:

Praktische Vorgehensweise:

- Zunächst Normalverteilung mit der gesuchten Fläche skizzieren
- Dann entsprechende Werte aus z-Tabelle auswählen

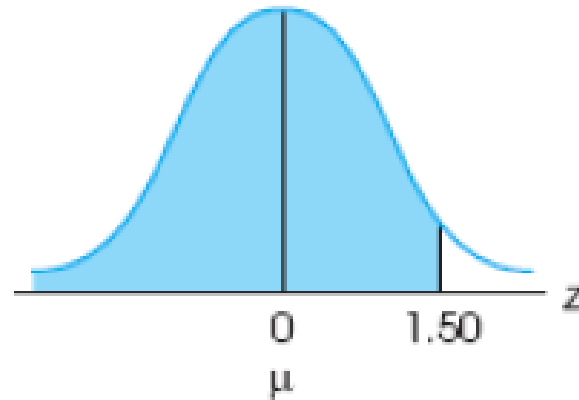


## Übungsbeispiel 2)



Welcher Flächenanteil der Normalverteilung entspricht  $z$ -Werten  $< 1,5$ ?  
Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen  $z$ -Wert  $< 1.5$  zu erhalten?

## Übungsbeispiel 2)



Welcher Flächenanteil der Normalverteilung entspricht z-Werten  $< 1,5$ ?  
Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen z-Wert  $< 1.5$  zu erhalten?

Vorgehen:

- Skizzieren NV und gesuchte Fläche
- Bestimme  $z = 1.5$  in der z-Werte Tabelle:

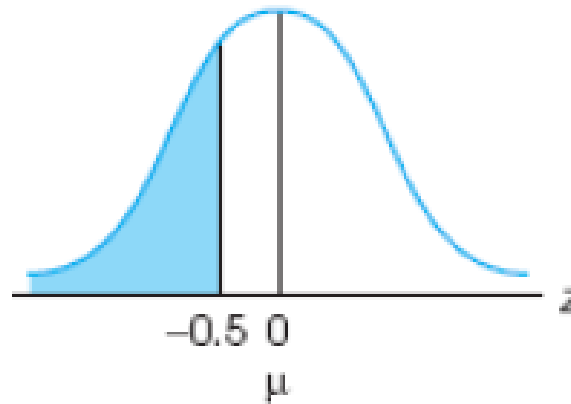
$$P(z < 1.5) = 0.9332 = 93.32\%$$

## Übungsbeispiel 3)

Welcher Flächenanteil der Normalverteilung entspricht z-Werten  $< -0,5$ ?

Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen z-Wert  $< -0,5$  zu erhalten?

## Übungsbeispiel 3): Hausaufgabe bzw. Tutorium



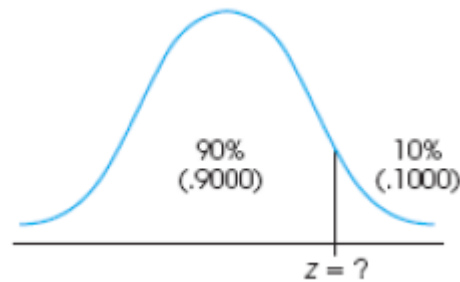
Welcher Flächenanteil der Normalverteilung entspricht  $z$ -Werten  $< -0,5$ ?

Wie groß ist die Wahrscheinlichkeit, für normalverteilte Werte einen  $z$ -Wert  $< -0,5$  zu erhalten?

## Übungsbeispiel 4)

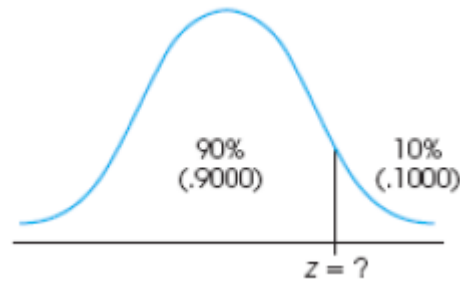
Welcher z-Wert separiert die obersten 10% aller Werte von den restlichen 90% der Verteilung?

## Übungsbeispiel 4)



Welcher z-Wert separiert die obersten 10% aller Werte von den restlichen 90% der Verteilung?

## Übungsbeispiel 4)



Welcher z-Wert separiert die obersten 10% aller Werte von den restlichen 90% der Verteilung?

- Skizzieren der Normalverteilung und der gesuchten Fläche
- Bestimme  $P = 0.90$  in der z-Werte Tabelle
- Bestimme korrespondierenden z-Wert:  $z = 1.28$

## Übungsbeispiel 5): Hausaufgabe bzw. Tutorium



Welche z-Werte separieren die mittleren 60% aller Werte von den restlichen 40% der Verteilung?



# Flächenanteile & Wahrscheinlichkeiten für z-Werte

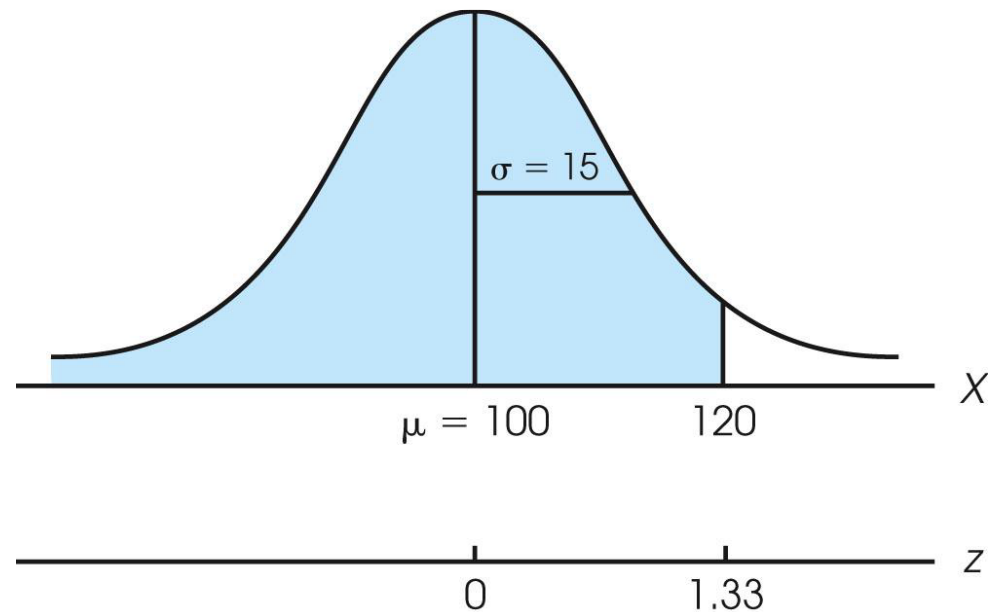
## Anwendungsbeispiel A:

- Gegeben sei eine Verteilung von IQ-Werten mit  $\mu = 100$  und  $\sigma = 15$ . Wie groß ist die Wahrscheinlichkeit, zufällig eine Person mit einem  $\text{IQ} < 120$  auszuwählen?

# Flächenanteile & Wahrscheinlichkeiten für z-Werte

## Anwendungsbeispiel A:

- Gegeben sei eine Verteilung von IQ-Werten mit  $\mu = 100$  und  $\sigma = 15$ . Wie groß ist die Wahrscheinlichkeit, zufällig eine Person mit einem  $\text{IQ} < 120$  auszuwählen?

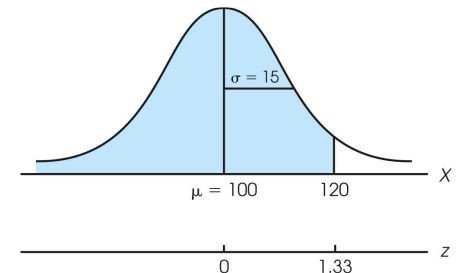


# Flächenanteile & Wahrscheinlichkeiten für z-Werte

- Anwendungsbeispiel A:
- Gegeben sei eine Verteilung von IQ-Werten mit  $\mu = 100$  und  $\sigma = 15$ . Wie groß ist die Wahrscheinlichkeit, zufällig eine Person mit einem  $\text{IQ} < 120$  auszuwählen?

## 1) Transformieren Rohwerte in z-Werte

$$z = \frac{X - \mu}{\sigma} = \frac{120 - 100}{15} = \frac{20}{15} = 1.33$$



IQ-Wert von 120 entspricht einem z-Wert von 1.33

IQ-Werte kleiner als 120 entsprechen z-Werten kleiner als 1.33

2) Korrespondierenden z-Wert in Tabelle auswählen:

$P = 0.9082$

**$P(X < 120) = P(z < 1.33) = 0.9082 = 90.82\%$**

# Flächenanteile & Wahrscheinlichkeiten für z-Werte

## Anwendungsbeispiel B:

- Wahrscheinlichkeiten bzw. Anteile zwischen zwei (normalverteilten) X-Werten bestimmen
- In der Gießener Innenstadt werden Geschwindigkeitsmessungen für Autofahrer durchgeführt. Bei der letzten Überprüfung sei für Autofahrer eine Durchschnitts-Geschwindigkeit von  $\mu = 58 \text{ km/h}$  mit einer Standardabweichung von  $\sigma = 10$  festgestellt worden. Die Messwerte seien (näherungsweise) normalverteilt.
- Wie hoch ist der Anteil der Autofahrer, die zwischen  $55 \text{ km/h}$  und  $65 \text{ km/h}$  in der Gießener Innenstadt fahren?

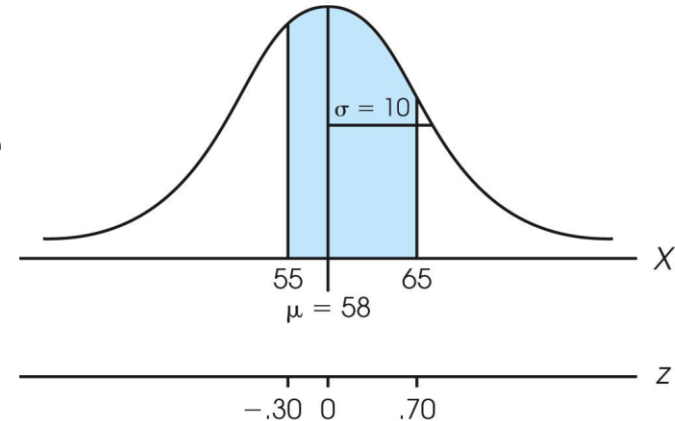
# Flächenanteile & Wahrscheinlichkeiten für z-Werte

## Anwendungsbeispiel B:

### 1) Transformieren der Rohwerte in z-Werte

$$\text{Für } X = 55\text{km/h: } z = \frac{X - \mu}{\sigma} = \frac{55 - 58}{10} = -\frac{3}{10} = -0.3$$

$$\text{Für } X = 65\text{km/h: } z = \frac{X - \mu}{\sigma} = \frac{65 - 58}{10} = \frac{7}{10} = 0.7$$



### 2. Verteilung mit gesuchtem Intervall skizzieren

#### 3a. Bestimmen der Fläche links von $X = 65$

$$\text{Für } z = .70, P = 0.76$$

#### 3b. Bestimmen der Fläche links von $X = 55$

$$\text{Für } z = -.30, P = 0.38$$

$$4. \text{ Subtrahieren: } 0.76 - 0.38 = 0.38$$

## Anwendungsbeispiel B:

- Wahrscheinlichkeiten/Anteile zwischen zwei (normalverteilten) X-Werten bestimmen
- In der Gießener Innenstadt werden Geschwindigkeitsmessungen für Autofahrer durchgeführt. Bei der letzten Überprüfung sei für Autofahrer eine Durchschnitts-Geschwindigkeit von  $\mu = 58 \text{ km/h}$  mit einer Standardabweichung von  $\sigma = 10$  festgestellt worden. Die Messwerte seien (näherungsweise) normalverteilt.
- Wie hoch ist der Anteil der Autofahrer, die zwischen  $55 \text{ km/h}$  und  $65 \text{ km/h}$  in der Gießener Innenstadt fahren? → 38%

# Flächenanteile & Wahrscheinlichkeiten für z-Werte

## Anwendungsbeispiel C:

### X-Werte für Wahrscheinlichkeiten/Anteile bestimmen

- Der Asta der JLU finanziert eine sozialwissenschaftliche Untersuchung zur Belastung durch Pendeln unter Studierenden. Die Ergebnisse zeigen, dass von den Studierenden im Durchschnitt  $\mu = 24.3$  Minuten pro Studientag für An-und Abreise verbraucht werden; die Standardabweichung sei  $\sigma = 10$ .
- Wieviel Minuten müssten Sie mindestens pendeln, um zu den 10% Studis mit der höchsten Pendeldauer für An-und Abreise zum Studienort zu gehören?

## Flächenanteile & Wahrscheinlichkeiten für z-Werte

- Anwendungsbeispiel C: X-Werte für Wahrscheinlichkeiten/Anteile bestimmen
  1. Bestimme 90% bzw. 0.90 in der z-Werte Tabelle und den dazugehörigen z-Wert:  $z = 1.282$
  2. Bestimme das Vorzeichen des gesuchten z-Wertes: positiv
  3. Transformiere den z-Wert in den Rohwert:

$$\begin{aligned} X &= \mu + z\sigma \\ &= 24.3 + 1.282 \cdot 10 \\ &= 24.3 + 12.82 \\ &= 37.1 \end{aligned}$$



## Flächenanteile & Wahrscheinlichkeiten für z-Werte

### Anwendungsbeispiel C:

#### X-Werte für Wahrscheinlichkeiten/Anteile bestimmen

- Der Asta der JLU finanziert eine sozialwissenschaftliche Untersuchung zur Belastung durch Pendeln unter Studierenden. Die Ergebnisse zeigen, dass von den Studierenden im Durchschnitt  $\mu = 24.3$  Minuten pro Studientag für An-und Abreise verbraucht werden; die Standardabweichung sei  $\sigma = 10$ .
- Wieviel Minuten müssten Sie mindestens pendeln, um zu den 10% Studis mit der höchsten Pendeldauer für An-und Abreise zum Studienort zu gehören?

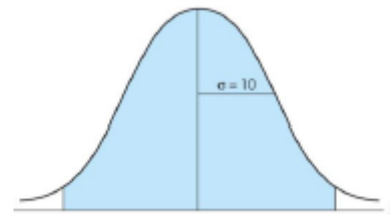
→ ca. 37 Minuten

# Flächenanteile & Wahrscheinlichkeiten für z-Werte

## Hausaufgabe / Tutorium!

Anwendungsbeispiel D (gleiche Population wie eben):

- X-Werte zwischen zwei Wahrscheinlichkeiten/Anteilswerten bestimmen
- Wie lautet die Spannweite für die mittleren 90% der Verteilung?



- Dichtefunktion der Normalverteilung als Hilfsmittel, um Häufigkeiten bzw. Wahrscheinlichkeiten für kontinuierliche Variablen zu ermitteln
- Wahrscheinlichkeiten können als (Flächen-)Anteile interpretiert werden
- Für normalverteilte Daten liegen tabellarische Darstellungen für interessierende Anteilwerte/Wahrscheinlichkeiten vor, die mit den jeweiligen z-Werten korrespondieren
  - Anhand der Formel zur z-Transformation können X-Werte in z-Werte und z-Werte in X-Werte transformiert werden
  - Für z-Werte können die zugehörigen Wahrscheinlichkeiten/Anteile aus der z-Tabelle entnommen werden

# Lernziele

- Sie erweitern Ihre Kenntnisse über die sog. „Normalverteilung“ und wissen wozu sie in der Statistik dient
- Sie können Flächenanteile und damit Wahrscheinlichkeiten innerhalb der Normalverteilung berechnen