5. Vorlesung Statistik I
Wahrscheinlichkeitstheorie II

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.

Erwartungswert und Varianz von Zufallsvariablen

Motivation

- Genau wie für die „normalen“ Variablen aus der Deskriptivstatistik lassen sich auch für Zufallsvariablen Lage- und Streuungsmaße definieren:
  - Lagemaße zur Beschreibung von einer „typischen“ Realisation der ZV.
  - Streuungsmaße zur Beschreibung des Ausmaßes der Unterschiedlichkeit der Realisationen.

- Wir werden die folgenden Maßzahlen besprechen:
  - als wichtigstes Lagemaß den **Erwartungswert**.
  - als wichtigste Streuungsmaße die **Varianz** und die **Standardabweichung**.

Erwartungswert einer Zufallsvariablen

Wiederholung: Definition des Mittelwerts

- Zur Erinnerung: In der Deskriptivstatistik haben wir als Lagemaß den Mittelwert kennengelernt.
- Gegeben seien die Messwerte $x_1, x_2, \dots, x_n$. Der Mittelwert ist definiert als

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

- Eine äquivalente Definition ist:

$$\bar{x} = \sum_{j=1}^{m} x_j \cdot h(x_j)$$

- Das heißt, jeder Summand entspricht einer Messwertausprägung x_j multipliziert mit der relativen Häufigkeit $h(x_j)$.
- Wie könnte sich diese zweite Formel auf diskrete Zufallsvariablen X übertragen lassen?

Definition: Erwartungswert einer diskreten ZV

- Lösung: Gleiche Formel, aber
  - statt Messwertausprägungen Realisationen
  - statt relativen Häufigkeiten Wahrscheinlichkeiten

- Der **Erwartungswert** $E(X)$ einer diskreten ZV *X* ist also:

$$E(X) = \sum_{j=1}^{m} x_j \cdot P\big(X = x_j\big)$$

- Der Erwartungswert $E(X)$ einer diskreten Zufallsvariablen X entspricht somit der Summe aller möglichen Realisationen jeweils multipliziert mit der Wahrscheinlichkeit.

- Falls eine Wahrscheinlichkeitsfunktion *f* von *X* gegeben ist, gilt für den Erwartungswert wegen $f\big(x_j\big) = P\big(X = x_j\big)$:

$$E(X) = \sum_{j=1}^{m} x_j \cdot P\big(X = x_j\big) = \sum_{j=1}^{m} x_j \cdot f\big(x_j\big)$$

Definition: Erwartungswert einer stetigen ZV

- Um den Erwartungswert für stetige ZVs zu definieren, gehen wir von der Formel für diskrete ZVs aus, ersetzen aber
  - die Summe durch ein Integral
  - die Wahrscheinlichkeitsfunktion durch eine Wahrscheinlichkeitsdichtefunktion

- Der **Erwartungswert** $E(X)$ einer stetigen ZV *X* ist also:

$$E(X) = \int_{-\infty}^{+\infty} x \cdot f(x) dx$$

Interpretation des Erwartungswerts

- Der Erwartungswert $E(X)$ einer Zufallsvariablen X ist eine konkrete Zahl.

- Er kann sowohl für diskrete als auch für stetige Zufallsvariablen wie folgt interpretiert werden:

- Der Erwartungswert $E(X)$ einer Zufallsvariable *X* entspricht der **durchschnittlichen Realisation von *X*,** wenn das X zugrundeliegende Zufallsexperiment unendlich oft wiederholt wird.

Beispiel: Berechnung des Erwartungswerts

- In einer Lostrommel befinden sich 10 Lose, darunter ein Gewinn zu 5 Euro und zwei Gewinne zu je 2 Euro. Ein Los kostet 1 Euro.
- Die Zufallsvariable *X* stehe für den erzielten Gewinn bzw. Verlust.
- $T_x = \{-1, 1, 4\}$ mit $x_1 = -1$, $x_2 = 1$ und $x_3 = 4$.
- Wahrscheinlichkeitsfunktion von *X*:

x_j	$f(x_j)$
-1	0.7
1	0.2
4	0.1

- Erwartungswert von *X*:

$$E(X) = \sum_{j=1}^{3} x_j \cdot f(x_j) = -1 \cdot 0.7 + 1 \cdot 0.2 + 4 \cdot 0.1 = -0.1$$

- Interpretation?

Rechenregeln für den Erwartungswert

- Seien X und Y und X_1, X_2, …, X_n diskrete oder stetige ZVs und *a* eine beliebige reelle Zahl (eine Konstante). Dann gilt:

$$E(a) = a$$

$$E(X + a) = E(X) + a$$

$$E(a \cdot X) = a \cdot E(X)$$

$$E(X + Y) = E(X) + E(Y)$$

$$E\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} E(X_i)$$

Zwischengliederung

- Bislang:
  - Erwartungswert

- Jetzt:
  - Varianz und Standardabweichung

Varianz und Standardabweichung einer Zufallsvariablen

Wiederholung: Definition der empirischen Varianz

- Zur Erinnerung: In der Deskriptivstatistik haben wir als Streuungsmaß die empirische Varianz kennengelernt.
- Gegeben seien die Messwerte $x_1, x_2, \dots, x_n$. Die empirische Varianz ist definiert als:

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

- Eine äquivalente Definition ist:

$$s_{emp}^2 = \sum_{j=1}^{m}(x_j - \bar{x})^2 \cdot h(x_j)$$

- Das heißt, jeder Summand entspricht der quadratischen Abweichung einer Messwertausprägung x_j vom arithmetischen Mittel $\bar{x}$ multipliziert mit der relativen Häufigkeit $h(x_j)$.
- Wie könnte sich dies auf diskrete Zufallsvariablen X übertragen lassen?

Definition: Varianz einer diskreten ZV

- Lösung: Gleiche Formel, aber
  - statt Messwertausprägungen Realisationen
  - statt relativen Häufigkeiten Wahrscheinlichkeiten
  - statt Mittelwert Erwartungswert

- Die **Varianz** $Var(X)$ einer diskreten ZV *X* ist also:

$$Var(X) = \sum_{j=1}^{m} (x_j - E(X))^2 \cdot P\big(X = x_j\big)$$

- Falls eine Wahrscheinlichkeitsfunktion *f* von *X* gegeben ist, gilt für die Varianz wegen $f\big(x_j\big) = P\big(X = x_j\big)$:

$$Var(X) = \sum_{j=1}^{m} (x_j - E(X))^2 \cdot P\big(X = x_j\big) = \sum_{j=1}^{m} (x_j - E(X))^2 \cdot f\big(x_j\big)$$

Definition: Varianz einer stetigen ZV

- Um die Varianz für stetige ZVs zu definieren, gehen wir von der Formel für diskrete ZVs aus, ersetzen aber
  - die Summe durch ein Integral
  - die Wahrscheinlichkeitsfunktion durch eine Wahrscheinlichkeitsdichtefunktion

- Die **Varianz** $Var(X)$ einer stetigen ZV *X* ist also:

$$Var(X) = \int_{-\infty}^{+\infty} \left(x - E(X)\right)^2 \cdot f(x)dx$$

Definition: Standardabweichung einer ZV

- Die **Standardabweichung** $SD(X)$ ist sowohl für diskrete als auch für stetige ZVs definiert als die Wurzel der Varianz:

$$SD(X) = \sqrt{Var(X)}$$

Interpretation der Varianz und der Standardabweichung

- Die Varianz $Var(X)$ und die Standardabweichung $SD(X)$ einer Zufallsvariablen X sind konkrete Zahlen.

- Sie können sowohl für diskrete als auch für stetige Zufallsvariablen wie folgt interpretiert werden:

- Die Varianz $Var(X)$ einer Zufallsvariable *X* entspricht der **empirischen Varianz der Realisationen von *X*,** wenn das X zugrundeliegende Zufallsexperiment unendlich oft wiederholt wird.

- Die Standardabweichung $SD(X)$ einer Zufallsvariable *X* entspricht der **empirischen Standardabweichung der Realisationen von *X*,** wenn das X zugrundeliegende Zufallsexperiment unendlich oft wiederholt wird.

- Also: Genau wie bei empirischer Varianz und Standardabweichung nicht wirklich intuitiv.

Beispiel: Berechnung der Varianz und der Standardabweichung

- Losbeispiel von Folie 9. Zur Erinnerung: $E(X) = -0.1$

x_j	$x_j - E(X)$	$\left(x_j - E(X)\right)^2$	$f(x_j)$
-1	- 0.9	0.81	0.7
1	1.1	1.21	0.2
4	4.1	16.81	0.1

$$Var(X) = \sum_{j=1}^{3} (x_j - E(X))^2 \cdot f\left(x_j\right) = 0.81 \cdot 0.7 + 1.21 \cdot 0.2 + 16.81 \cdot 0.1 = 2.49$$

$$SD(X) = \sqrt{Var(X)} = \sqrt{2.49} = 1.58$$

Rechenregeln für die Varianz und die Standardabweichung

- Sei X eine diskrete oder stetige ZV und *a* eine beliebige reelle Zahl. Dann gilt:

$$Var(X + a) = Var(X)$$

$$SD(X + a) = SD(X)$$

$$Var(a \cdot X) = a^2 \cdot Var(X)$$

$$SD(a \cdot X) = a \cdot SD(X)$$

Weitere Maße zur Beschreibung von Zufallsvariablen

- Neben dem Erwartungswert und der Varianz gibt es viele weitere Maßzahlen zur Beschreibung von Zufallsvariablen.

- Beispielsweise lassen sich auch Median, Modus und Interquartilsabstand von Zufallsvariablen definieren. Diese werden wir hier aber nicht besprechen.

- Außerdem ist es möglich, für zwei Zufallsvariablen X und Y Kovarianz und Korrelation zu definieren, um deren Zusammenhang zu beschreiben. Damit werden wir uns im Rahmen der Vorlesung Diagnostik I / Statistik III beschäftigen.

Zwischengliederung

- Bislang:
  - Erwartungswert
  - Varianz und Standardabweichung

- Jetzt:
  - z-Standardisierung von Zufallsvariablen

z-Standardisierung von Zufallsvariablen

z-Standardisierung von Zufallsvariablen I

- Zur Erinnerung: In der Deskriptivstatistik haben wir Variablen z-standardisiert, indem wir von jedem Messwert den Mittelwert abgezogen haben und diese Differenz durch die empirische Standardabweichung geteilt haben:

$$z_i = \frac{x_i - \bar{x}}{s_{emp\ x}}$$

- Für z-standardisierte Variablen gilt:
  - $\bar{z} = 0$
  - $s_{emp\ z} = 1.$

z-Standardisierung von Zufallsvariablen II

- Auch Zufallsvariablen können z-standardisiert werden. Hierfür ziehen wir von der Zufallsvariable *X* ihren Erwartungswert $E(X)$ ab und teilen diese Differenz durch die Standardabweichung $SD(X)$ von *X*. Diese Transformation ergibt eine neue **z-standardisierte Zufallsvariable *Z***:

$$Z = \frac{X - E(X)}{SD(X)}$$

- Für z-standardisierte Zufallsvariablen *Z* gilt:
  - $\mathrm{E}(Z) = 0$
  - $SD(Z) = 1.$

Zwischengliederung

- Bislang:
  - Erwartungswert
  - Varianz und Standardabweichung
  - z-Standardisierung von Zufallsvariablen

- Jetzt:
  - Konkrete Wahrscheinlichkeitsverteilungen

Konkrete Wahrscheinlichkeitsverteilungen von Zufallsvariablen

Konkrete Wahrscheinlichkeitsverteilungen I

- Zur Erinnerung: Eine Wahrscheinlichkeitsverteilung P_X einer Zufallsvariablen *X* weist jedem Ereignis A_X der Zufallsvariable eine Wahrscheinlichkeit $P_X(A_X)$ zu.

- Wenn wir also eine konkrete Wahrscheinlichkeitsverteilung P_X angeben wollen, müssen wir für jedes Ereignis A_X eine konkrete Wahrscheinlichkeit angeben.

- Problem: Dies ist in den meisten Fällen extrem aufwendig bis unmöglich.

- Lösung: Charakterisierung von P_X durch Angabe einer
  - Wahrscheinlichkeitsfunktion im Falle einer diskreten ZV *X*.
  - Wahrscheinlichkeitsdichtefunktion im Falle einer stetigen ZV *X*.

- Dann können die Wahrscheinlichkeiten $P_X(A_X)$ aller Ereignisse A_X einfach mithilfe dieser berechnet werden.

Konkrete Wahrscheinlichkeitsverteilungen II

- Wir werden daher konkrete Wahrscheinlichkeitsverteilungen jeweils durch Angabe ihrer Wahrscheinlichkeitsfunktion oder ihrer Wahrscheinlichkeitsdichte definieren.

- Die folgenden Wahrscheinlichkeitsverteilungen werden wir kennenlernen:
    - **Bernoulli-Verteilung**
    - **Binomialverteilung**
    - **Normalverteilung**

- Für alle diese Wahrscheinlichkeitsverteilungen werden wir jeweils
    - die Wahrscheinlichkeitsfunktion bzw. Wahrscheinlichkeitsdichtefunktion
    - die Verteilungsfunktion
    - den Erwartungswert, die Varianz und die Standardabweichung betrachten.

Bernoulli-Verteilung

Wahrscheinlichkeitsfunktion I

- Als **Bernoulli-Variable** bezeichnet man eine Zufallsvariable X, die sich nur in den Werten $x_1 = 0$ und $x_2 = 1$ realisieren kann.

- Eine Bernoulli-Variable ist also eine diskrete ZV mit Träger $T_x = \{0,1\}$.

- Häufig wird die Realisation 1 als „Erfolg“ und die Realisation 0 als „Misserfolg“ bezeichnet.

- Beispiel 1: Beim Wurf eines sechsseitigen Würfels wird den Augenzahlen 1, 2 und 3 durch die ZV *X* der Wert 1 und den Augenzahlen 4, 5 und 6 der Wert 0 zugeordnet.

- Beispiel 2: Bei einer zufällig gewählten Person wird durch die ZV X der Merkmalsausprägung „braune Augen“ der Wert 1 und der Merkmalsausprägung „andere Augenfarbe“ der Wert 0 zugeordnet.

Wahrscheinlichkeitsfunktion II

- Um die Wahrscheinlichkeitsfunktion einer Bernoulli-Variablen anzugeben, müssen wir die Wahrscheinlichkeiten aller Elementarereignisse angeben, also die Wahrscheinlichkeit für „Erfolg“ $P(X = 1)$ und die Wahrscheinlichkeit für „Misserfolg“ $P(X = 0)$.

- Sei also $P(X = 1) = \pi$, wobei π ein beliebiger Wert zwischen 0 und 1 ist. Dann ist $P(X = 0) = 1 - P(X = 1) = 1 - \pi$

- Die Wahrscheinlichkeitsfunktion einer Bernoulli-Variable dann:

$$f(0) = P(X = 0) = 1 - \pi$$

$$f(1) = P(X = 1) = \pi$$

Hinweis:
Das Symbol π steht hier NICHT für die Zahl „Pi“ ≈ 3.14

- Beziehungsweise etwas kompakter mit $x_1 = 0$ und $x_2 = 1$:

$$f(x_j) = \pi^{x_j}(1 - \pi)^{1-x_j}$$

Hinweis: $a^0 = 1$ für beliebiges a

- Die zu dieser Wahrscheinlichkeitsfunktion gehörende Wahrscheinlichkeitsverteilung wird **Bernoulli-Verteilung** genannt.

Wahrscheinlichkeitsfunktion III

- Die konkrete Wahrscheinlichkeitsfunktion einer Bernoulli-Variable hängt davon ab, welchem konkreten Wert π entspricht.

- Für jeden konkreten Wert π ergibt sich eine andere Wahrscheinlichkeitsfunktion und somit auch eine andere konkrete Bernoulli-Verteilung.

- Zum Beispiel:
  - π = 0.2: $f(x_j) = 0.2^{x_j} \cdot 0.8^{1-x_j}$
  - π = 0.5: $f(x_j) = 0.5^{x_j} \cdot 0.5^{1-x_j} = 0.5$
  - π = 0.7: $f(x_j) = 0.7^{x_j} \cdot 0.3^{1-x_j}$
  - usw.

Wahrscheinlichkeitsfunktion IV

- Eine Größe, deren Ausprägung die konkrete Form der Wahrscheinlichkeitsverteilung innerhalb einer Klasse von Wahrscheinlichkeitsverteilungen festlegt, nennt man **Parameter**.

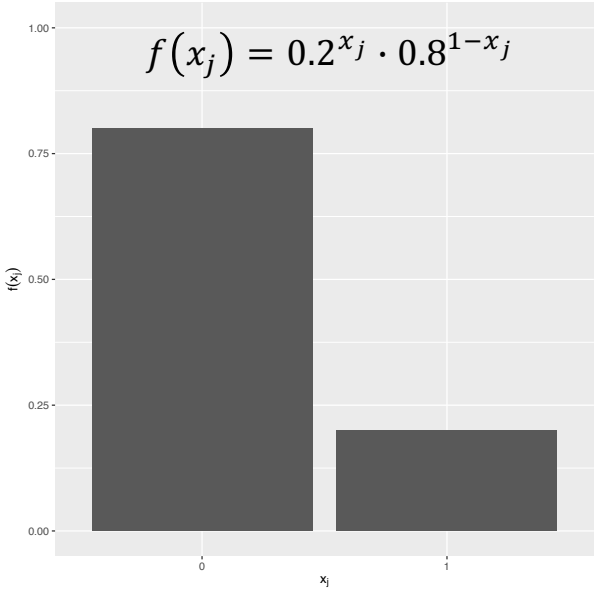
- π ist der Parameter der Klasse der Bernoulli-Verteilungen.

- Wenn wir sagen wollen, dass eine ZV *X* einer Bernoulli-Verteilung mit Parameter π folgt, schreiben wir: $X \sim Be(\pi)$

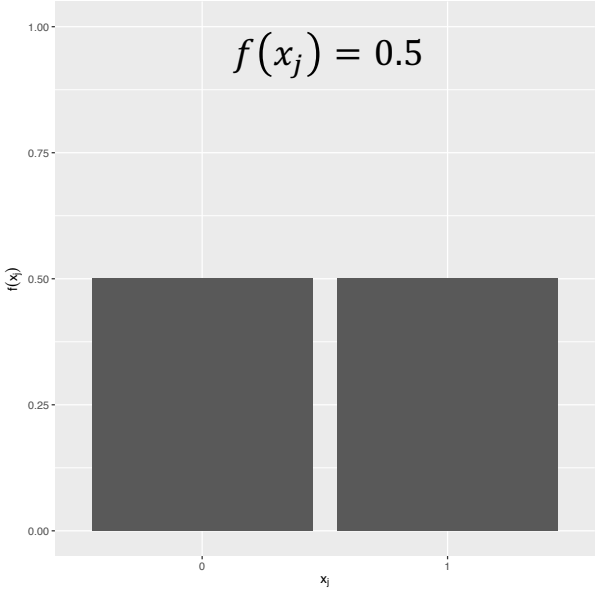
Wahrscheinlichkeitsfunktion V

- Graphische Darstellung der Wahrscheinlichkeitsfunktionen einiger ZVs mit Bernoulli-Verteilung und unterschiedlichen Werten für den Parameter π:

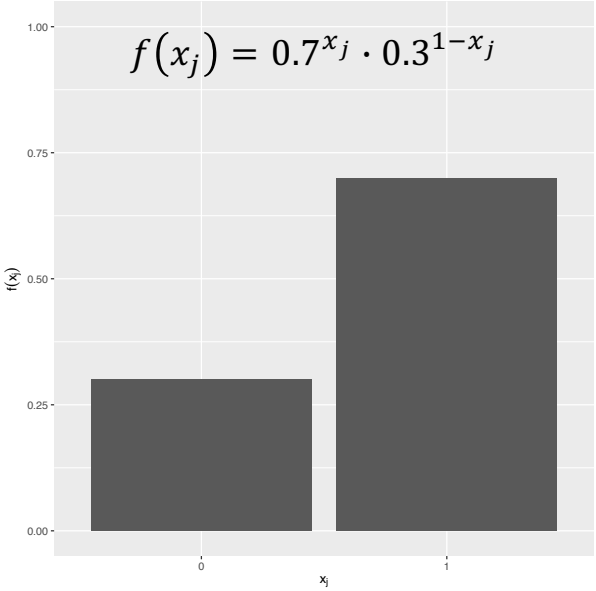
π = 0.2



π = 0.5



π = 0.7



- Interpretation π: Wahrscheinlichkeit dafür, dass sich die Bernoulli-Variable im Wert 1 realisiert, also Wahrscheinlichkeit für „Erfolg“.

Verteilungsfunktion I

- Zur Erinnerung: Bei diskreten Zufallsvariablen besteht folgender Zusammenhang zwischen der Wahrscheinlichkeitsfunktion und der Verteilungsfunktion:

$$F(x_k) = \sum_{j=1}^{k} f\left(x_j\right)$$

Verteilungsfunktion II

- Für ZVs mit Bernoulli-Verteilung ist die Verteilungsfunktion also:

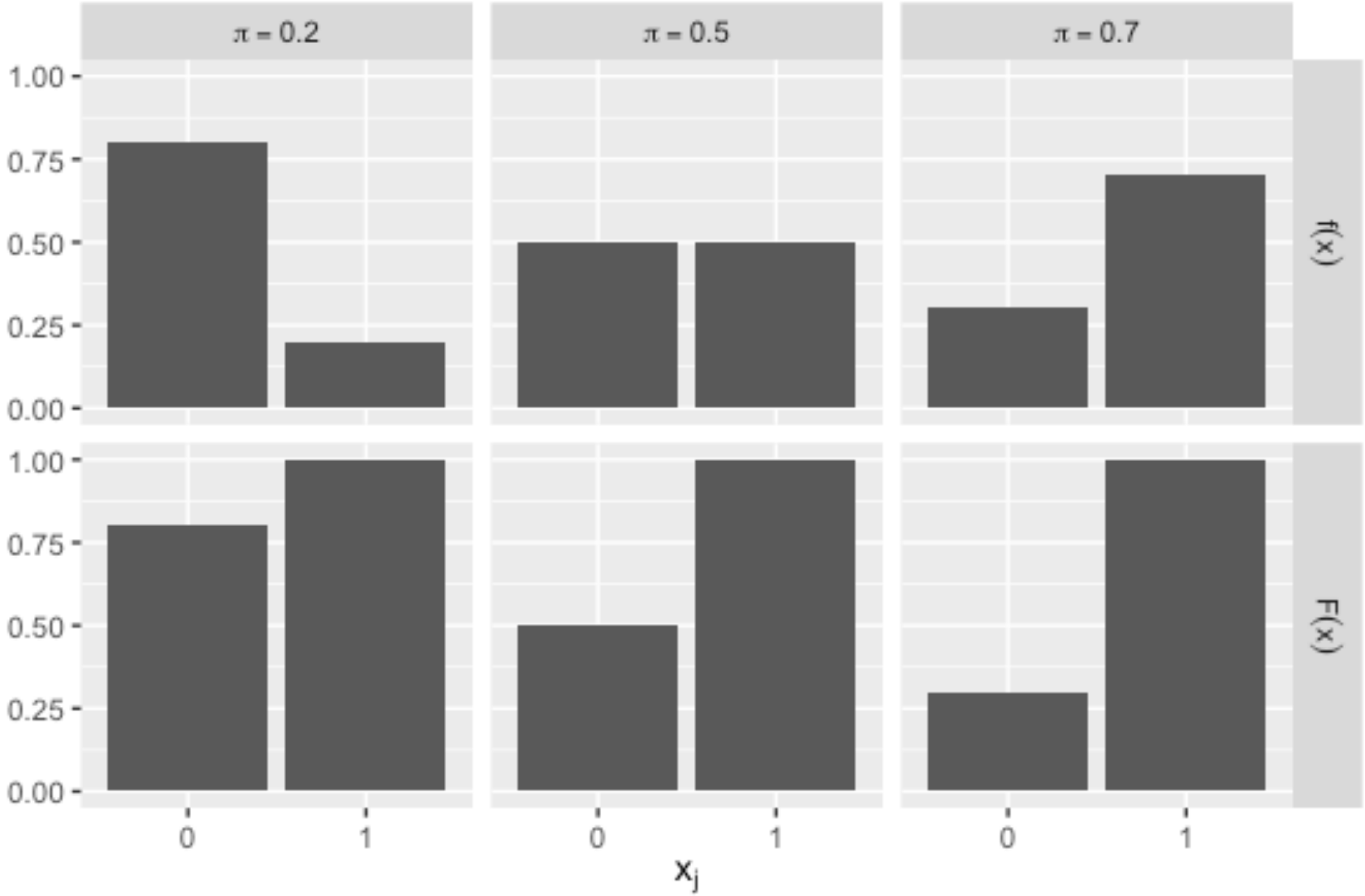
$$F(x_k) = \sum_{j=1}^{k} f\left(x_j\right) = \sum_{j=1}^{k} \pi^{x_j}(1-\pi)^{1-x_j}$$

- Einsetzen von $x_1 = 0$ und $x_2 = 1$:

$$F(0) = F(x_1) = \sum_{j=1}^{1} f\left(x_j\right) = f(x_1) = 1 - \pi$$

$$F(1) = F(x_2) = \sum_{j=1}^{2} f\left(x_j\right) = f(x_1) + f(x_2) = \pi + (1-\pi) = 1$$

Verteilungsfunktion III



Erwartungswert, Varianz und Standardabweichung I

- Der Erwartungswert einer ZV mit Bernoulli-Verteilung mit Parameter π kann leicht berechnet werden:

$$E(X) = \sum_{j=1}^{m} x_j \cdot f(x_j) = x_1 \cdot f(x_1) + x_2 \cdot f(x_2) = 0 \cdot (1-\pi) + 1 \cdot \pi = \pi$$

- Daraus folgt für die Varianz:

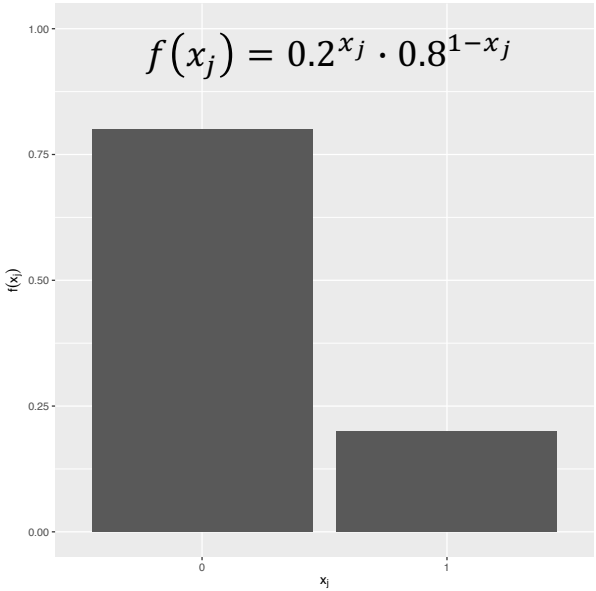
$$Var(X) = \sum_{j=1}^{m} \left(x_j - E(X)\right)^2 \cdot f(x_j) = (0-\pi)^2 \cdot (1-\pi) + (1-\pi)^2 \cdot \pi$$
$$= \pi^2 - \pi^3 + \pi - 2\pi^2 + \pi^3 = \pi(1-\pi)$$

- Die Standardabweichung ist somit:

$$SD(X) = \sqrt{Var(X)} = \sqrt{\pi(1-\pi)}$$

Erwartungswert, Varianz und Standardabweichung II

π = 0.2

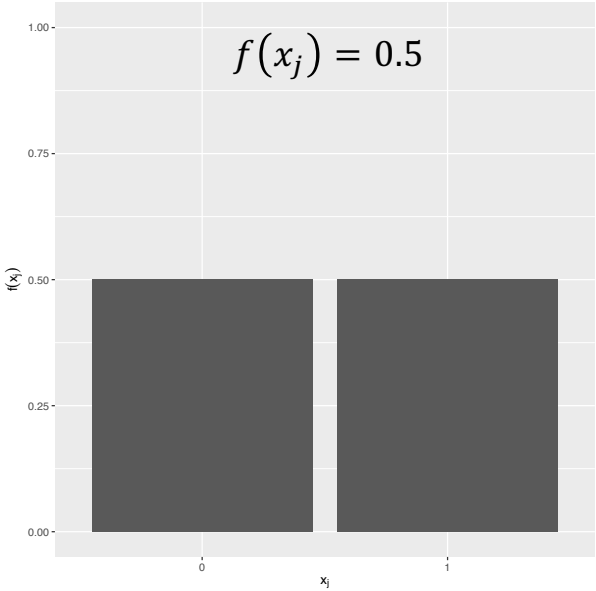


$E(X) = \pi = 0.2$

$Var(X) = \pi(1-\pi) = 0.16$

$SD(X) = \sqrt{Var(X)} = 0.4$

π = 0.5

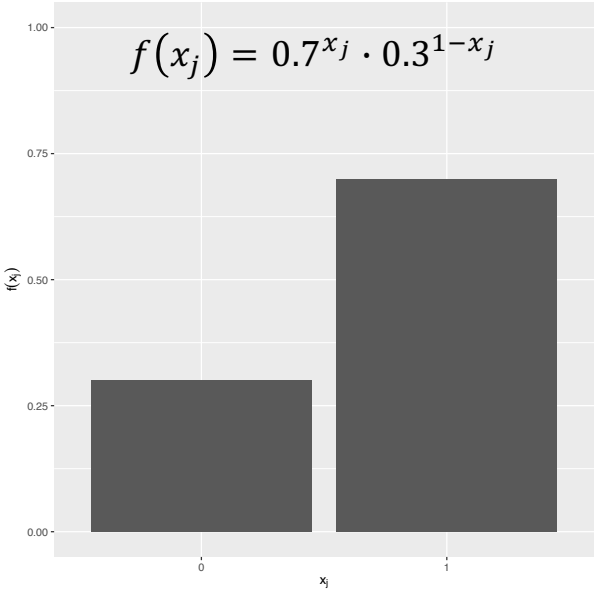


$E(X) = 0.5$

$Var(X) = 0.25$

$SD(X) = 0.5$

π = 0.7



$E(X) = 0.7$

$Var(X) = 0.21$

$SD(X) \approx 0.46$

- Weitere Interpretation von π (neben $\pi = P(X = 1)$): Erwartungswert von X

Binomialverteilung

Vorbereitungen

- Zunächst einige Vorbereitungen:
  - Binomialkoeffizient
  - Unabhängigkeit von Zufallsvariablen
  - Summen von unabhängigen Bernoulli-Variablen

Binomialkoeffizient I

- Definition der **Fakultät** einer natürlichen Zahl $n \in \mathbb{N}$:
  - Unter der Fakultät $n!$ versteht man das Produkt der ganzen Zahlen von 1 bis n: $n! = 1 \cdot 2 \cdot \dots \cdot n$
  - Per Definition ist $0! = 1$
  - Beispiele:

$$3! = 1 \cdot 2 \cdot 3 = 6$$

$$5! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 = 120$$

$$10! = 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \cdot \ \dots \cdot 10 = 3628800$$

Binomialkoeffizient II

- Definition des **Binomialkoeffizienten** für $n, k \in \mathbb{N}$ und $n > k$:
  - Der Ausdruck $\binom{n}{k} = \frac{n!}{k!(n-k)!}$ heißt Binomialkoeffizient
  - Beispiele:

$$\binom{5}{2} = \frac{5!}{2!\,(5-2)!} = \frac{120}{2! \cdot 3!} = \frac{120}{12} = 10$$

$$\binom{3}{0} = \frac{3!}{0!\,(3-0)!} = \frac{6}{6} = 1$$

Unabhängigkeit von Zufallsvariablen

- Die genaue Definition der Unabhängigkeit von Zufallsvariablen ist etwas schwierig und für uns nur selten praktisch relevant.

- Wir begnügen uns daher mit einer intuitiven Definition:

  Zwei Zufallsvariablen *X* und *Y* sind **unabhängig**, falls die Realisation der ZV *X* keine Auswirkungen darauf hat, wie sich die ZV *Y* realisiert und umgekehrt.

- Für zwei unabhängige Zufallsvariablen *X* und *Y* gibt es eine weitere Varianz-Rechenregel:

$$Var(X \pm Y) = Var(X) + Var(Y)$$

- Allgemein gilt für *n* unabhängige Zufallsvariablen $X_1, X_2, \ldots, X_n$:

$$Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i)$$

Exkurs: Unabhängigkeit von Ereignissen

- Die Ereignisse $A_1, \ldots, A_n$ sind unabhängig, wenn gilt:

$$P(A_1 \cap \cdots \cap A_n) = P(A_1) \cdot \cdots \cdot P(A_n)$$

Exkurs: Unabhängigkeit von Zufallsvariablen

- Unabhängigkeit von **diskreten Zufallsvariablen**:

- Die diskreten Zufallsvariablen $X_1, \dots, X_n$ sind unabhängig, wenn für beliebige Realisationen $x_1, \dots, x_n$ aus den jeweiligen Trägern der Zufallsvariablen gilt:

- $P(X_1 = x_1, \dots, X_n = x_n) = P(X_1 = x_1) \cdot \cdots \cdot P(X_n = x_n)$

- Unabhängigkeit von **stetigen Zufallsvariablen**:

- Die stetigen Zufallsvariablen $X_1, \dots, X_n$ sind unabhängig, wenn für beliebige Realisationen $x_1, \dots, x_n$ aus den jeweiligen Trägern der Zufallsvariablen gilt:

- $P(X_1 \leq x_1, \dots, X_n \leq x_n) = P(X_1 \leq x_1) \cdot \cdots \cdot P(X_n \leq x_n)$

Hinweis:
Die Schreibweise $P(X_1 = x_1, X_2 = x_2)$ bezeichnet die gemeinsame Wahrscheinlichkeit, dass die ZV X_1 den Wert x_1 annimmt und gleichzeitig die ZV X_2 den Wert x_2 annimmt.

Summen von unabhängigen Bernoulli-Variablen I

- Wir betrachten *n* unabhängige Bernoulli-Variablen $X_1, X_2, \ldots, X_n$ mit

$$X_i \sim Be(\pi)$$

für alle *i* = 1, 2, …, *n*

- Das heißt, alle *n* Bernoulli-Variablen sind unabhängig und folgen jeweils einer Bernoulli-Verteilung mit demselben Parameter π.

- Beispiel:
  - Zufallsexperiment*:* Eine Münze wird *n*-mal unabhängig voneinander geworfen.
  - X_i gibt jeweils an, ob beim *i*-ten Wurf Kopf oder Zahl oben landet:

$$X_i = \begin{cases} 0 & falls\ Zahl\ bei\ Wurf\ i\ oben\ landet \\ 1 & falls\ Kopf\ bei\ Wurf\ i\ oben\ landet \end{cases}$$

Summen von unabhängigen Bernoulli-Variablen II

- Die Summe dieser *n* Bernoulli-Variablen

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

  ist wieder eine Zufallsvariable.

- Da die einzelnen Bernoulli-Variablen X_i sich jeweils nur in den Werten 0 und 1 realisieren können, steht die ZV *X* für die Anzahl der Bernoulli-Variablen, die sich in dem Wert 1 realisiert haben, d.h. für die „Anzahl der Erfolge“.

- Im Münzwurf-Beispiel steht *X* für die Anzahl Kopf bei *n* Münzwürfen.

Summen von unabhängigen Bernoulli-Variablen III

- *X* kann ganzzahlige Werte zwischen 0 („kein Erfolg“) und *n* („*n* Erfolge“) annehmen, je nachdem wie viele der *n* einzelnen Bernoulli-Variablen sich in dem Wert 1 realisieren.

- Der Träger von *X* ist also $T_X = \{0, 1, 2, \dots, n\}$ und *X* ist somit eine diskrete ZV.

- Man kann beweisen, dass *X* die Wahrscheinlichkeitsfunktion

$$f(x_j) = \binom{n}{x_j} \pi^{x_j} (1-\pi)^{n-x_j}$$

besitzt, wobei $x_1 = 0$, $x_2 = 1$, … , $x_m = n$ ist.

- Dies ist die Wahrscheinlichkeitsfunktion der Binomialverteilung.

Wahrscheinlichkeitsfunktion I

- Eine Zufallsvariable X folgt einer **Binomialverteilung**, falls ihr Träger $T_X = \{0, 1, 2, \dots, n\}$ ist, und ihre Wahrscheinlichkeitsfunktion die Form

$$f(x_j) = \binom{n}{x_j} \pi^{x_j} (1 - \pi)^{n - x_j}$$

aufweist, wobei $x_1 = 0$, $x_2 = 1$, …, $x_m = n$ ist.

- Die Binomialverteilung hat zwei Parameter:
  - $n \in \mathbb{N}$.
  - $\pi \in [0, 1]$

- Wenn wir sagen wollen, dass eine ZV *X* einer Binomialverteilung mit Parametern n und π folgt, schreiben wir: $X \sim B(n, \pi)$

Wahrscheinlichkeitsfunktion II

- Konkrete Beispiele für unterschiedliche Werte der Parameter *n* und π:

  - $n = 10$ und $\pi = 0.2$:

$$f(x_j) = \binom{n}{x_j} \pi^{x_j} (1-\pi)^{n-x_j} = \binom{10}{x_j} 0.2^{x_j} (1-0.2)^{10-x_j} = \binom{10}{x_j} 0.2^{x_j} \cdot 0.8^{10-x_j}$$

  - $n = 10$ und $\pi = 0.5$:

$$f(x_j) = \binom{n}{x_j} \pi^{x_j} (1-\pi)^{n-x_j} = \binom{10}{x_j} 0.5^{x_j} (1-0.5)^{10-x_j} = \binom{10}{x_j} 0.5^{10}$$

  - $n = 10$ und $\pi = 0.7$:

$$f(x_j) = \binom{n}{x_j} \pi^{x_j} (1-\pi)^{n-x_j} = \binom{10}{x_j} 0.7^{x_j} (1-0.7)^{10-x_j} = \binom{10}{x_j} 0.7^{x_j} \cdot 0.3^{10-x_j}$$

Wahrscheinlichkeitsfunktion III

- Beispiel für die Berechnung der Funktionswerte $f(x_j)$ für $n = 3$ und $\pi = 0.5$:

$$f(x_j) = \binom{n}{x_j} \pi^{x_j} (1-\pi)^{n-x_j}$$

$$f(x_1) = f(0) = \binom{3}{0} 0.5^0 (1-0.5)^{3-0} = 1 \cdot 0.5^0 \cdot 0.5^3 = 0.125$$

$$f(x_2) = f(1) = \binom{3}{1} 0.5^1 (1-0.5)^{3-1} = 3 \cdot 0.5^1 \cdot 0.5^2 = 0.375$$

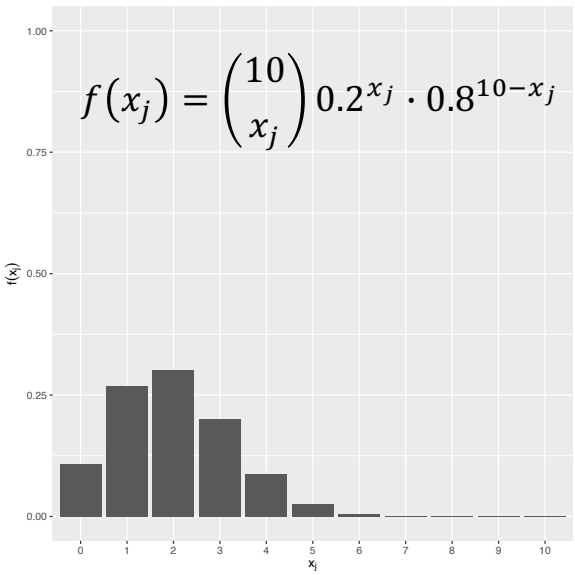
$$f(x_3) = f(2) = \binom{3}{2} 0.5^2 (1-0.5)^{3-2} = 3 \cdot 0.5^2 \cdot 0.5^1 = 0.375$$

$$f(x_4) = f(3) = \binom{3}{3} 0.5^3 (1-0.5)^{3-3} = 1 \cdot 0.5^3 \cdot 0.5^0 = 0.125$$

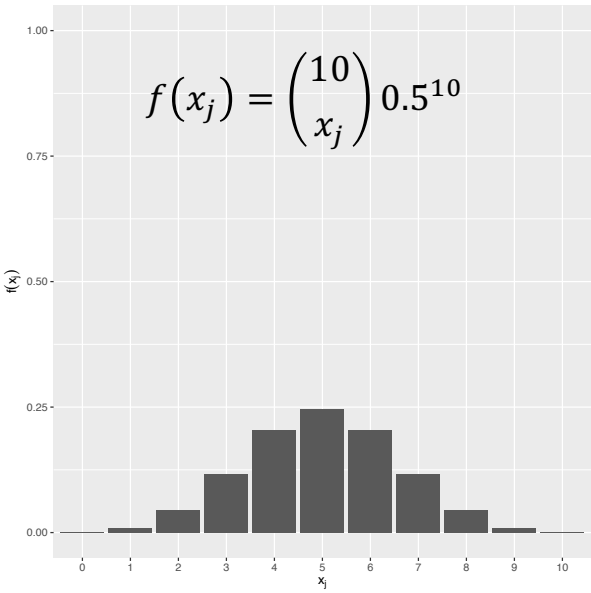
Wahrscheinlichkeitsfunktion IV

- Graphische Darstellung der Wahrscheinlichkeitsfunktionen einiger ZVs mit Binomialverteilung mit *n* = 10 und unterschiedlichen Werten für den Parameter π:

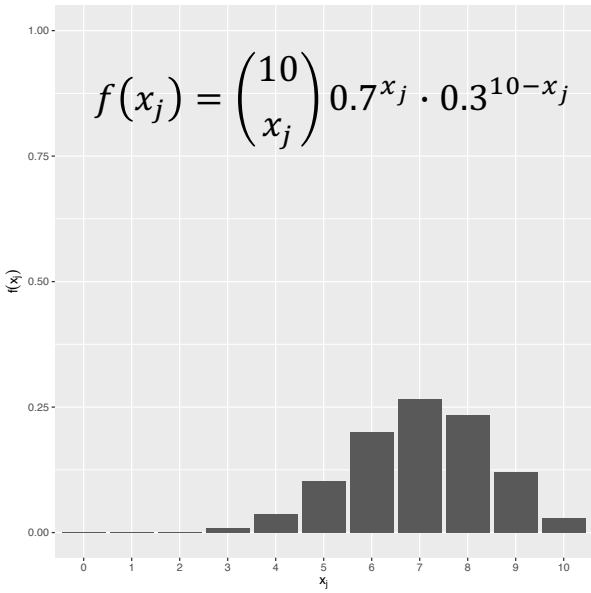
n = 10
π = 0.2



n = 10
π = 0.5



n = 10
π = 0.7



- Interpretation *n*: Anzahl der zugrundeliegenden Bernoulli-Variablen X_i.
- Interpretation π: Wahrscheinlichkeit für Erfolg der zugrundeliegenden Bernoulli-Variablen X_i.

Verteilungsfunktion I

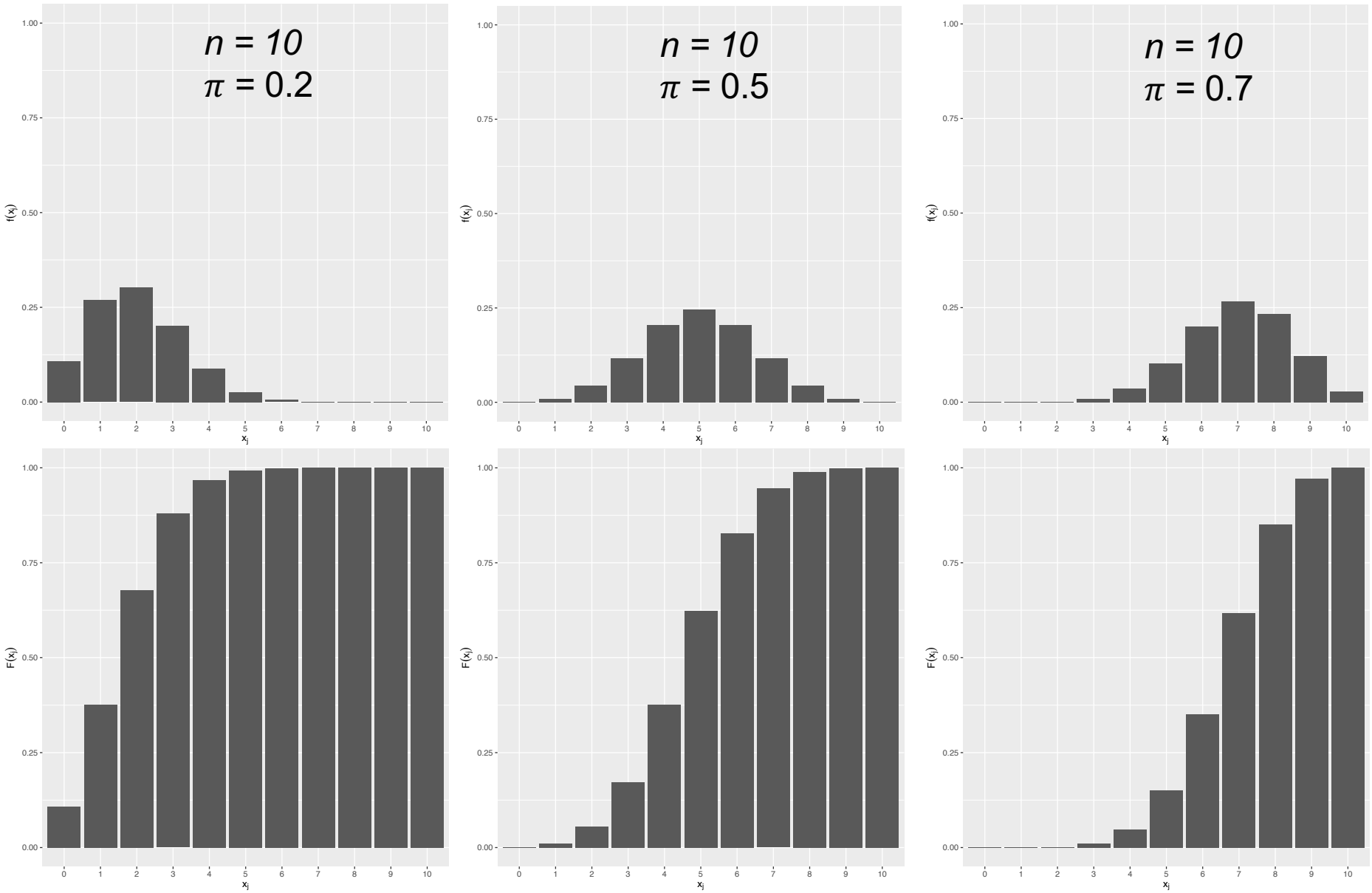
- Zur Erinnerung: Bei einer diskreten Zufallsvariable X gilt für die Verteilungsfunktion:

$$F(x_k) = P(X \leq x_k) = \sum_{j=1}^{k} f(x_j)$$

- Die Verteilungsfunktion einer binomialverteilten Zufallsvariable X ist somit:

$$F(x_k) = \sum_{j=1}^{k} f(x_j) = \sum_{j=1}^{k} \binom{n}{x_j} \pi^{x_j} (1-\pi)^{n-x_j}$$

Verteilungsfunktion II



Erwartungswert, Varianz und Standardabweichung I

- Der Erwartungswert der binomialverteilten Zufallsvariable X ergibt sich aus den Erwartungswerten der einzelnen Bernoulli-Variablen X_i :

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

- Die Varianz ergibt sich aus den Varianzen der einzelnen Bernoulli-Variablen X_i und deren Unabhängigkeit:

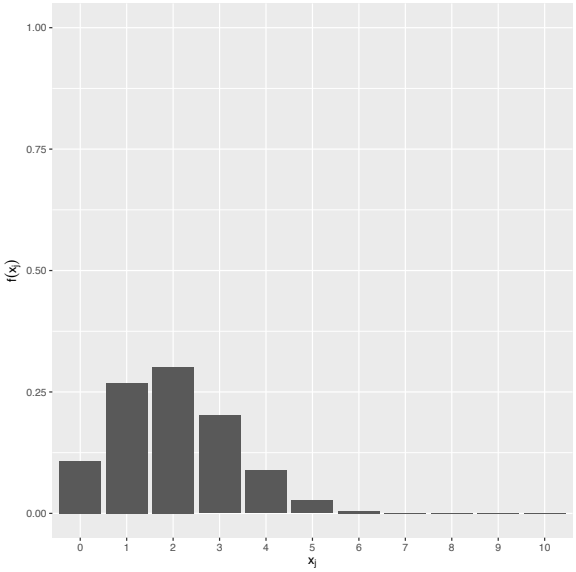
$$Var(X) = Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = \sum_{i=1}^{n} \pi(1-\pi) = n\pi(1-\pi)$$

- Standardabweichung:

$$SD(X) = \sqrt{Var(X)} = \sqrt{n\pi(1-\pi)}$$

Erwartungswert, Varianz und Standardabweichung II

n = 10
π = 0.2

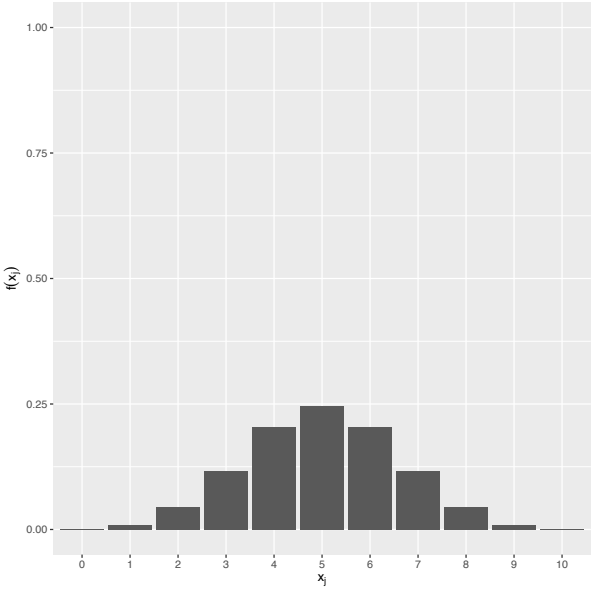


$$E(X) = n\pi = 2$$

$$Var(X) = n\pi(1-\pi) = 1.6$$

$$SD(X) = \sqrt{Var(X)} \approx 1.26$$

n = 10
π = 0.5

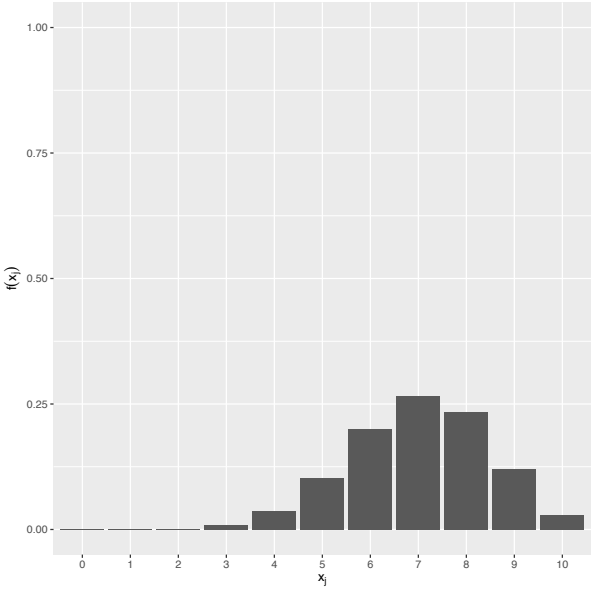


$$E(X) = 5$$

$$Var(X) = 2.5$$

$$SD(X) \approx 1.58$$

n = 10
π = 0.7



$$E(X) = 7$$

$$Var(X) = 2.1$$

$$SD(X) \approx 1.45$$

Beispiel I

- Zufallsexperiment: Faire Münze wird zehn Mal geworfen.
- Die Zufallsvariable *X* steht für die Anzahl Kopf.
- *X* folgt somit einer Binomialverteilung mit Parametern *n* = 10 und π = 0.5:

$$X \sim B(10, 0.5)$$

- Der Erwartungswert von *X* ist $E(X) = 5$.
- Interpretation: Würden wir unendlich oft zehn faire Münzen werfen, würde im Durchschnitt 5 mal Kopf oben landen.

Beispiel II

- Die Wahrscheinlichkeitsfunktion von *X* ist

$$f(x_j) = \binom{n}{x_j} \pi^{x_j} (1-\pi)^{n-x_j} = \binom{10}{x_j} 0.5^{10}$$

- Damit können wir die Wahrscheinlichkeiten für alle möglichen Ereignisse berechnen.

- Zum Beispiel:
  - A_{X1} „sieben mal Kopf“:

$$P_X(A_{X1}) = f(7) = \binom{10}{7} 0.5^{10} \approx 0.12$$

  - A_{X2} „zwei oder fünf mal Kopf“:

$$P_X(A_{X2}) = f(2) + f(5) = \binom{10}{2} 0.5^{10} + \binom{10}{5} 0.5^{10} \approx 0.29$$

  - A_{X3} „mindestens acht mal Kopf“:

$$P_X(A_{X3}) = f(8) + f(9) + f(10) = \binom{10}{8} 0.5^{10} + \binom{10}{9} 0.5^{10} + \binom{10}{10} 0.5^{10} \approx 0.05$$

Normalverteilung

Wahrscheinlichkeitsdichtefunktion I

- Eine stetige Zufallsvariable X folgt einer **Normalverteilung**, falls ihr Träger aus den gesamten reellen Zahlen besteht, also $T_X = \mathbb{R}$ ist, und ihre Wahrscheinlichkeitsdichte die Form

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right)$$

hat, wobei π hier für die Zahl „Pi“ (≈ 3.14) steht.

- Die Normalverteilung hat zwei Parameter:
  - $\mu \in \mathbb{R}$
  - $\sigma^2 \in \mathbb{R}_+$

- Wenn wir sagen wollen, dass eine ZV X einer Normalverteilung mit Parametern μ und σ^2 folgt, schreiben wir: $X \sim N(\mu, \sigma^2)$

Wahrscheinlichkeitsdichtefunktion II

- Konkrete Beispiele für unterschiedliche Werte der Parameter μ und σ^2:
  - $\mu = 1$ und $\sigma^2 = 1$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right) = \frac{1}{\sqrt{2\pi \cdot 1}} exp\left(-\frac{1}{2} \cdot \frac{(x-1)^2}{1}\right) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{(x-1)^2}{2}\right)$$

  - $\mu = 0$ und $\sigma^2 = 1$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right) = \frac{1}{\sqrt{2\pi \cdot 1}} exp\left(-\frac{1}{2} \cdot \frac{(x-0)^2}{1}\right) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right)$$

  - $\mu = 0$ und $\sigma^2 = 2$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right) = \frac{1}{\sqrt{2\pi \cdot 2}} exp\left(-\frac{1}{2} \cdot \frac{(x-0)^2}{2}\right) = \frac{1}{\sqrt{4\pi}} exp\left(-\frac{x^2}{4}\right)$$

Wahrscheinlichkeitsdichtefunktion III

- Beispiel für die Berechnung einiger Funktionswerte $f(x)$ für $\mu = 0$ und $\sigma^2 = 1$:

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{x^2}{2}\right)$$

$$f(-2) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{(-2)^2}{2}\right) = \frac{1}{\sqrt{2\pi}} exp(-2) = 0.05$$

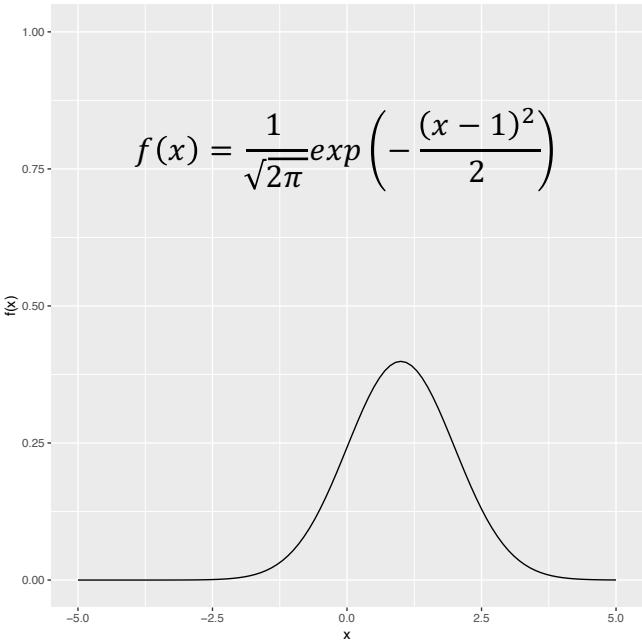
$$f(0) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{0^2}{2}\right) = \frac{1}{\sqrt{2\pi}} exp(0) = \frac{1}{\sqrt{2\pi}} \approx 0.40$$

$$f(1) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1^2}{2}\right) = \frac{1}{\sqrt{2\pi}} exp\left(-\frac{1}{2}\right) \approx 0.24$$

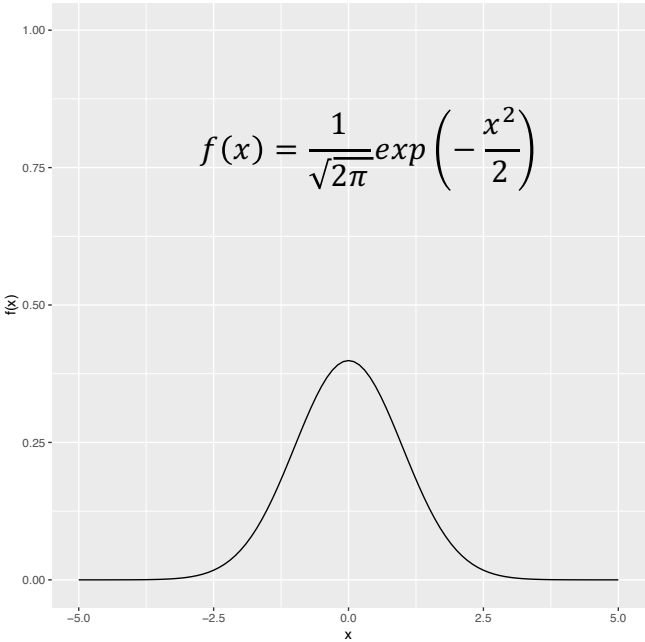
Wahrscheinlichkeitsdichtefunktion IV

- Graphische Darstellung der Wahrscheinlichkeitsdichtefunktionen einiger ZVs mit Normalverteilung und unterschiedlichen Werten für die Parameter μ und σ^2:

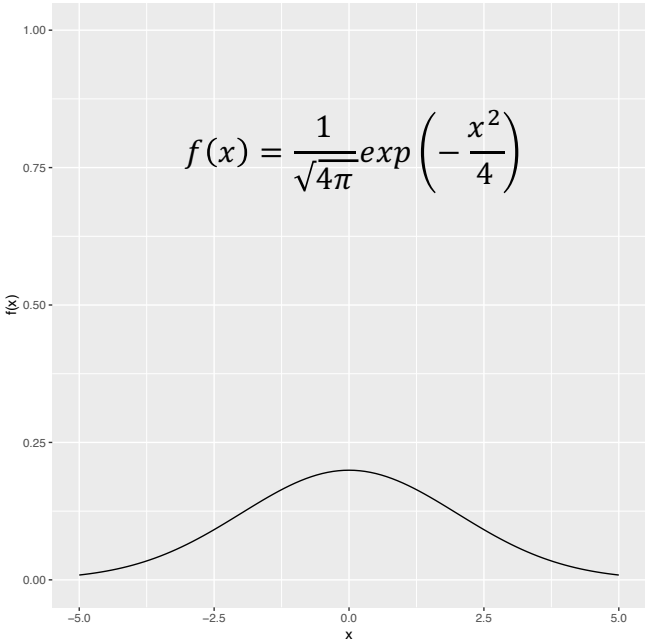
$\mu = 1$ und $\sigma^2 = 1$



$\mu = 0$ und $\sigma^2 = 1$



$\mu = 0$ und $\sigma^2 = 2$



- Der Parameter μ bestimmt, an welcher Stelle der x-Achse sich das Maximum der Dichtefunktion befindet.
- Der Parameter σ^2 legt die „Breite“ der Dichtefunktion fest.

Wahrscheinlichkeitsdichtefunktion V

- Wichtige Eigenschaften der Normalverteilung:

- Ihre Dichtefunktion $f(x)$ hat ihr Maximum an der Stelle $x = \mu$.

- Ihre Dichtefunktion $f(x)$ ist symmetrisch um μ, d.h. $f(\mu + c) = f(\mu - c)$ für alle $c \in \mathbb{R}$. Hieraus folgt unter anderem auch:

  - $P(X \leq \mu - c) = P(X \geq \mu + c)$
  - $P(X \leq \mu) = 0.5$

- Ihre Dichtefunktion hat eine „Glockenform“. Je weiter *x* von μ entfernt ist, desto kleiner ist die Dichte $f(x)$.

Verteilungsfunktion I

- Zur Erinnerung: Bei einer stetigen Zufallsvariable X gilt für die Verteilungsfunktion:

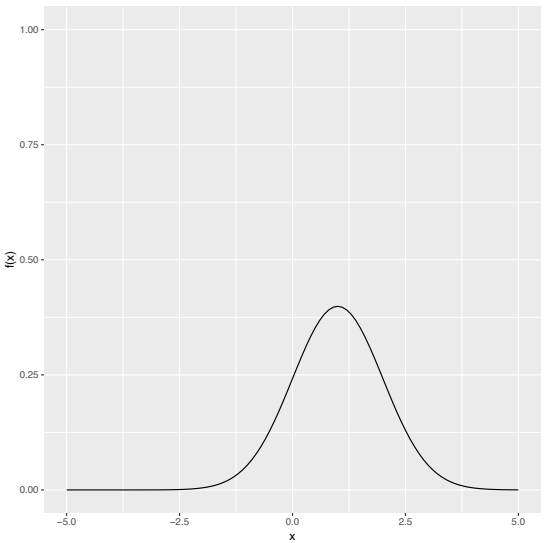
$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(y)dy$$

- Die Verteilungsfunktion einer normalverteilten Zufallsvariable X ist somit:

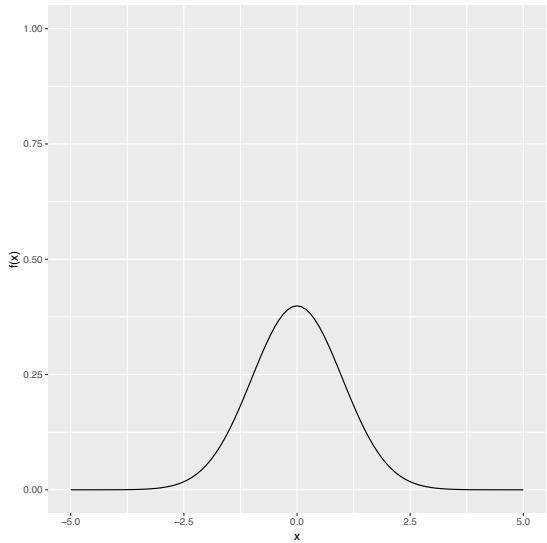
$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(y-\mu)^2}{\sigma^2}\right) dy$$

Verteilungsfunktion II

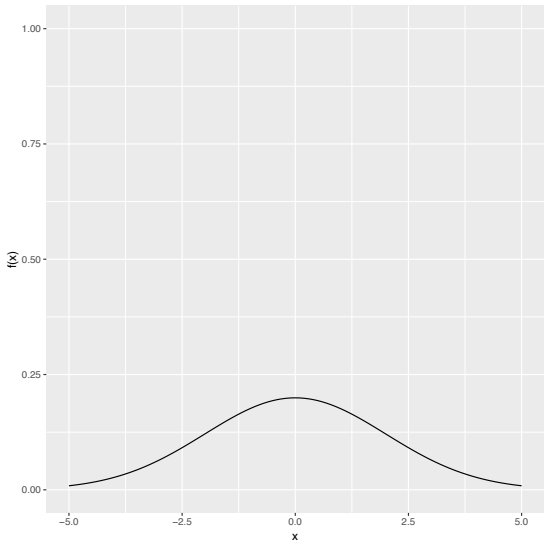
$\mu = 1$ und $\sigma^2 = 1$

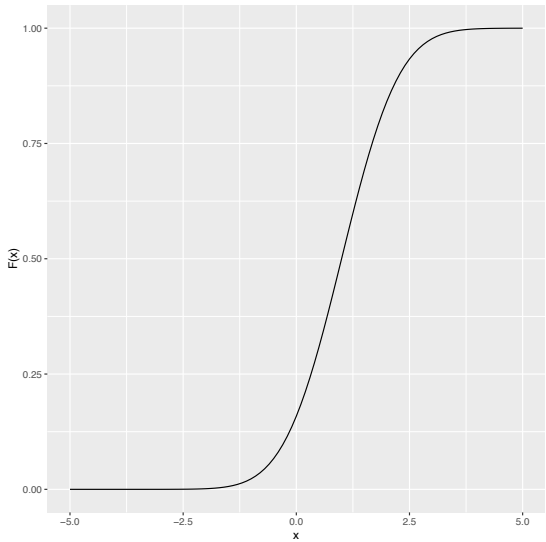


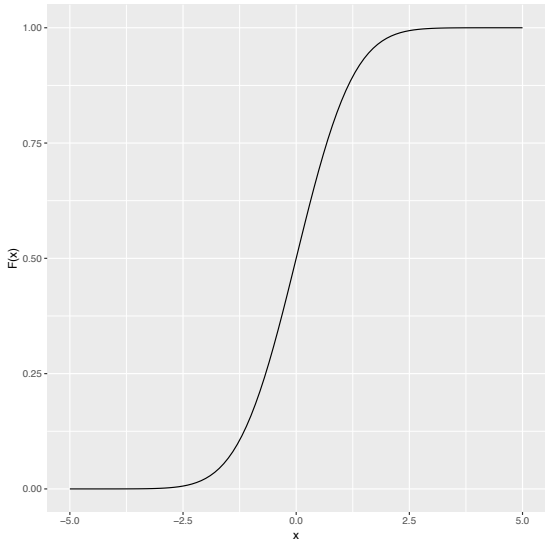
$\mu = 0$ und $\sigma^2 = 1$

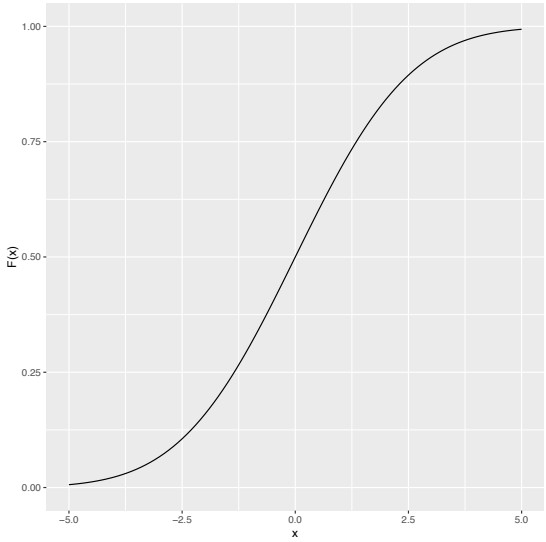


$\mu = 0$ und $\sigma^2 = 2$









Erwartungswert, Varianz und Standardabweichung I

- Der Erwartungswert einer normalverteilten Zufallsvariable *X* kann aus der allgemeinen Formel für stetige ZVs berechnet werden (sehr aufwendig):

$$\boldsymbol{E(X)} = \int_{-\infty}^{+\infty} x \cdot f(x)dx = \int_{-\infty}^{+\infty} x \cdot \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right) dx = \cdots = \boldsymbol{\mu}$$

- Ebenso die Varianz:

$$\boldsymbol{Var(X)} = \int_{-\infty}^{+\infty} \left(x - E(X)\right)^2 \cdot f(x)dx = \int_{-\infty}^{+\infty} (x-\mu)^2 \cdot \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2} \cdot \frac{(x-\mu)^2}{\sigma^2}\right) = \cdots = \boldsymbol{\sigma^2}$$

- Als Standardabweichung ergibt sich

$$SD(X) = \sqrt{Var(X)} = \sqrt{\sigma^2} = \sigma$$

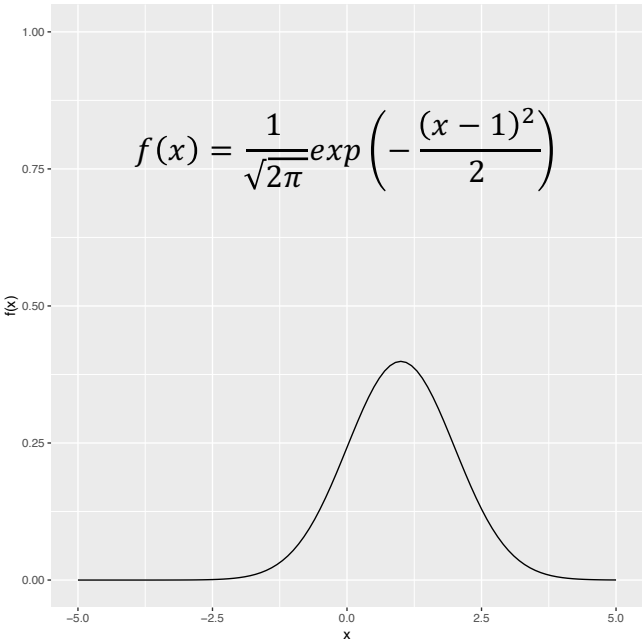
Erwartungswert, Varianz und Standardabweichung II

- Die Parameter μ und σ^2 entsprechen somit dem Erwartungswert und der Varianz der normalverteilten Zufallsvariable *X*.

- <u>Hinweis:</u>
  - Die Schreibweise $X \sim N(\mu, \sigma^2)$ bedeutet implizit eigentlich $X \sim N\big(E(X), Var(X)\big)$ wobei hier $E(X) = \mu$ und $Var(X) = \sigma^2$ gilt.
  - Sprachlich: „Die Zufallsvariable X ist normalverteilt mit Erwartungswert μ und Varianz σ^2."
  - Später werden wir kombinierte Zufallsvariablen betrachten, die auch normalverteilt sind, deren Erwartungswert und Varianz sich aber teilweise aus mehreren Größen zusammensetzen.

Erwartungswert, Varianz und Standardabweichung III

$\mu = 1$ und $\sigma^2 = 1$

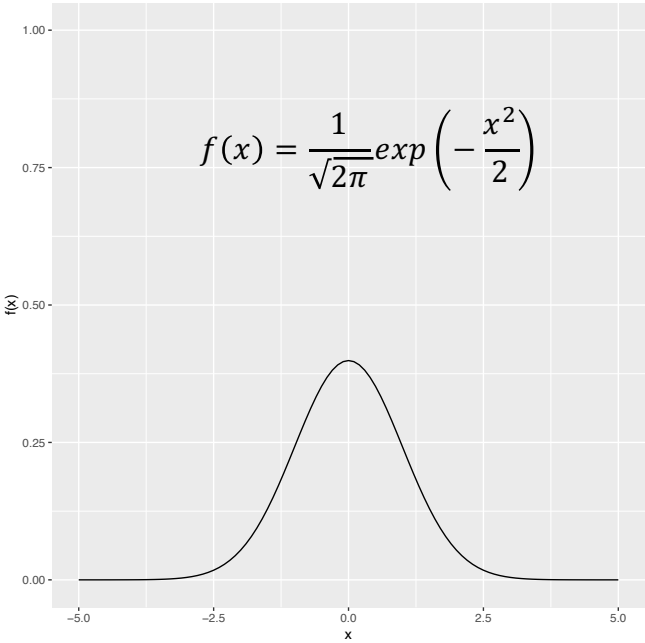


$E(X) = \mu = 1$

$Var(X) = \sigma^2 = 1$

$SD(X) = \sqrt{Var(X)} = 1$

$\mu = 0$ und $\sigma^2 = 1$

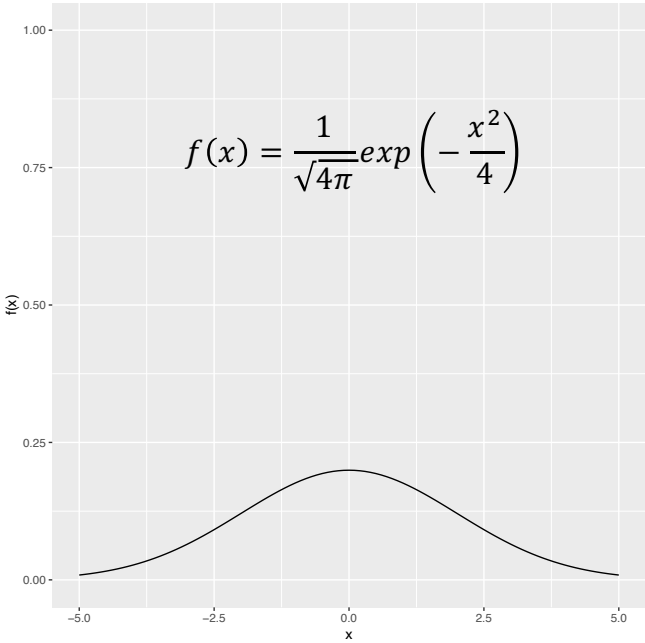


$E(X) = 0$

$Var(X) = 1$

$SD(X) = 1$

$\mu = 0$ und $\sigma^2 = 2$



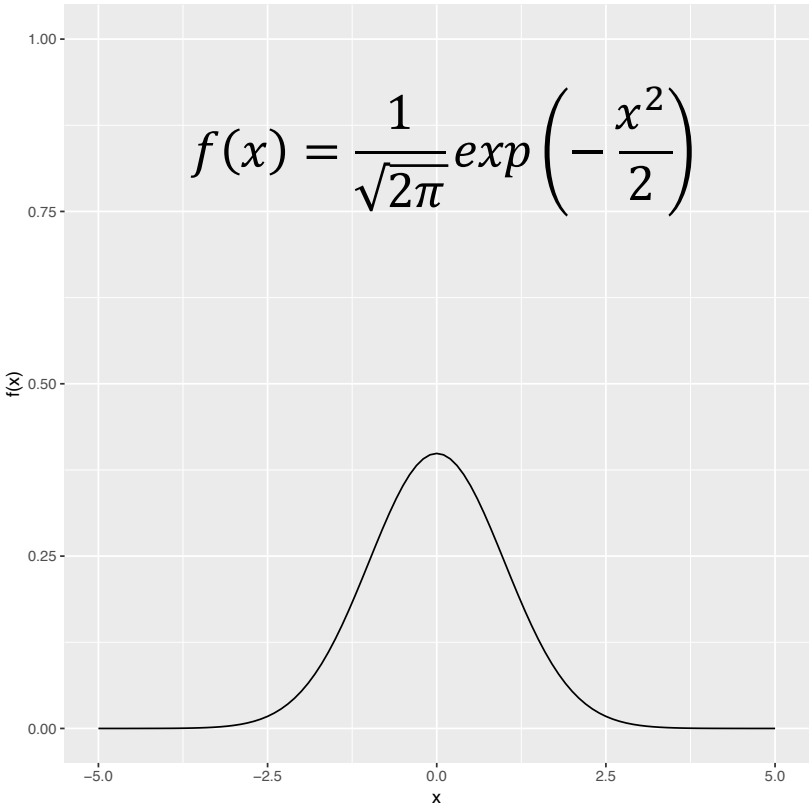
$E(X) = 0$

$Var(X) = 2$

$SD(X) = \sqrt{2} \approx 1.41$

Standardnormalverteilung

- Die Normalverteilung mit den Parameterwerten $\mu = 0$ und $\sigma^2 = 1$ hat einen eigenen Namen: Sie wird **Standardnormalverteilung** genannt.
- Der Erwartungswert einer standardnormalverteilten ZV *X* ist also $E(X) = \mu = 0$ und die Varianz ist $Var(X) = \sigma^2 = 1$
- Wahrscheinlichkeitsdichtefunktion der Standardnormalverteilung:



z-Standardisierung und Standardnormalverteilung

- Wendet man die z-Standardisierung auf eine beliebige normalverteilte Zufallsvariable X mit Parametern μ und σ^2 an, ergibt sich eine neue Zufallsvariable *Z*:

$$Z = \frac{X - E(X)}{SD(X)} = \frac{X - \mu}{\sigma}$$

- Diese neue Zufallsvariable *Z* folgt einer Standardnormalverteilung:

$$Z \sim N(0, 1)$$

  - dass $E(Z) = 0$ und $Var(Z) = 1$ ist, ergibt sich direkt aus der z-Standardisierung.
  - dass Z einer Normalverteilung folgt, ist etwas schwieriger zu beweisen.

- Dies bedeutet, dass jede normalverteilte Zufallsvariable *X* durch z-Standardisierung in eine standardnormalverteilte Zufallsvariable *Z* überführt werden kann.

- Dies ist eine sehr wichtige Eigenschaft von normalverteilten Zufallsvariablen, auf die wir in der Inferenzstatistik sehr häufig zurückgreifen werden.

Zwischengliederung

- Bislang:
  - Erwartungswert
  - Varianz und Standardabweichung
  - z-Standardisierung von Zufallsvariablen
  - Konkrete Wahrscheinlichkeitsverteilungen

- Jetzt:
  - Zusammenfassung

Zusammenfassung

- Zufallsvariablen können wie andere Variablen mit Maßzahlen beschrieben werden.
  - Analog zum arithmetischen Mittel ist der Erwartungswert $E(X)$ ein Maß der zentralen Tendenz.
  - Analog zur empirischen Varianz/Standardabweichung ist die Varianz $Var(X)$ bzw. Standardabweichung $SD(X)$ ein Maß der Unterschiedlichkeit (Streuung).
- Zufallsvariablen können wie andere Variablen standardisiert werden um eine neue ZV mit bekanntem Erwartungswert und bekannter Varianz zu erhalten.
- Für praktische Fragestellungen sind folgende konkrete Wahrscheinlichkeitsverteilungen in Zukunft für uns nützlich:

Verteilung	**Parameter**	**Träger**	***Wahrscheinlichkeitsfunktion bzw. Wahrscheinlichkeitsdichtefunktion***	$\boldsymbol{E(X)}$	$\boldsymbol{Var(X)}$
Bernoulli	$\pi \in [0,1]$	$T_X = \{0,1\}$	$f(x_j) = \pi^{x_j}(1-\pi)^{1-x_j}$	π	$\pi(1-\pi)$
Binomial	$n \in \mathbb{N}$ $\pi \in [0,1]$	$T_X = \{0,1,\dots,n\}$	$f(x_j) = \binom{n}{x_j}\pi^{x_j}(1-\pi)^{n-x_j}$	$n\pi$	$n\pi(1-\pi)$
Normal	$\mu \in \mathbb{R}$ $\sigma^2 \in \mathbb{R}_+$	$T_X = \mathbb{R}$	$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left(-\frac{1}{2}\cdot\frac{(x-\mu)^2}{\sigma^2}\right)$	μ	σ^2