1.5.4 Quantile und Modi

Bem. 1.73. [Quantil, Modus]

Analog zu Statistik I kann man auch Quantile und Modi definieren.
Gegeben sei eine Zufallsvariable X mit Wahrscheinlichkeitsverteilung P_X und Verteilungsfunktion F_X.

a) Für $\alpha \in (0,1)$ heißt x_α zugehöriges α-*Quantil* genau dann, wenn $F(x_\alpha) = \alpha$. Insbesondere ergeben sich für α=0.5 der *Median* und für α=0.25 bzw. 0.75 das *25%* bzw. *75% Quantil*.

b) Ist X diskret, so heißt jedes x_{mod} mit

$$P(X = x_{mod}) \geq P(X = x), \text{ für alle } x \in \mathcal{X}$$

Modus von P_X.

Ist X stetig mit Dichtefunktion f_X, so heißt jedes x_{mod} mit $f_X(x_{\text{mod}}) \geq f_X(x)$, für alle $x \in \mathcal{X}$ Modus von P_X (bezüglich der Dichte f_X).

Bem. 1.74. [Modalbereiche]

Ist der Modus x_{mod} eindeutig und X diskret, so ist dies derjenige Punkt, der die höchste Wahrscheinlichkeit besitzt, einzutreten.
Dies kann man auf “zusammenhängende“ Bereiche mit vorgegebener Sicherheit ausweiten:
Sei $\gamma \in (0,1)$. Ein Intervall $A_\gamma \subset \mathbb{R}$, $A_\gamma = [a,b]$, heiße *Modalintervall zum Niveau* γ, wenn gilt

$$P(X \in A_\gamma) \geq \gamma$$

und $P(X \in C) < \gamma$ für alle Intervalle $C \subset \mathbb{R}$, $C = [c,d]$ mit $d - c < b - a$.
Dann ist $A_\gamma \cap \mathcal{X}$ kleinster (“präzisester“) zusammengehöriger Bereich, der mindestens Wahrscheinlichkeit γ besitzt.
Ist X diskret, so ist $A_\gamma \cap \mathcal{X}$ ein Bereich benachbarter Punkte, bei stetigem X mit Träger $\mathbb{R}$ ist A_γ ein Intervall.

1.6 Wichtige Verteilungsmodelle

Wir behandeln hier nur die Binomial-, die Poisson- und die Normalverteilung. Einige weitere Verteilungsmodelle werden direkt dort eingeführt, wo sie benötigt werden.

1.6.1 Binomialverteilung

Konstruktionsprinzip:

- Ein Zufallsexperiment wird n mal unabhängig durchgeführt.
- Wir interessieren uns jeweils nur, ob ein bestimmtes Ereignis A eintritt oder nicht.
- $X =$ „absolute Häufigkeit, mit der Ereignis A bei n unabhängigen Versuchen eintritt“.
- Träger von X: $\mathcal{X} = \{0, 1, 2, \ldots, n\}$.

Herleitung der Wahrscheinlichkeitsfunktion:

- Bezeichne $\pi = P(A)$ die Wahrscheinlichkeit für A in $einem$ Experiment.
- Das Ereignis $\{X = x\}$ tritt z.B. auf, wenn in den ersten x Versuchen A eintritt und anschließend nicht mehr. Die Wahrscheinlichkeit dafür ist

$$\begin{aligned} P(A_1 \cap \ldots \cap A_x \cap \bar{A}_{x+1} \cap \ldots \cap \bar{A}_n) &= \underbrace{\pi \cdot \ldots \cdot \pi}_{x \text{ mal}} \cdot \underbrace{(1-\pi) \cdot \ldots \cdot (1-\pi)}_{n-x \text{ mal}} \\ &= \pi^x (1-\pi)^{n-x}. \end{aligned}$$

- Diese Überlegung gilt analog für jede Auswahl von x “Plätzen“. Insgesamt gibt es $\binom{n}{x}$ Möglichkeiten für die Verteilung der x Erfolge (Auftreten von A) auf n Plätze. Damit gilt:

$$P(X = x) = \binom{n}{x} \pi^x (1-\pi)^{n-x}.$$

Definition 1.75.

Eine Zufallsvariable heißt *binomialverteilt mit dem Parameter* π *bei* n *Versuchen*, kurz $X \sim B(n, \pi)$, wenn sie die Wahrscheinlichkeitsfunktion

$$f(x) = \begin{cases} \binom{n}{x} \pi^x (1-\pi)^{n-x}, & x = 0, 1, \ldots, n \\ 0, & \text{sonst} \end{cases}$$

besitzt.

Die $B(1, \pi)$-Verteilung heißt auch *Bernoulliverteilung*.

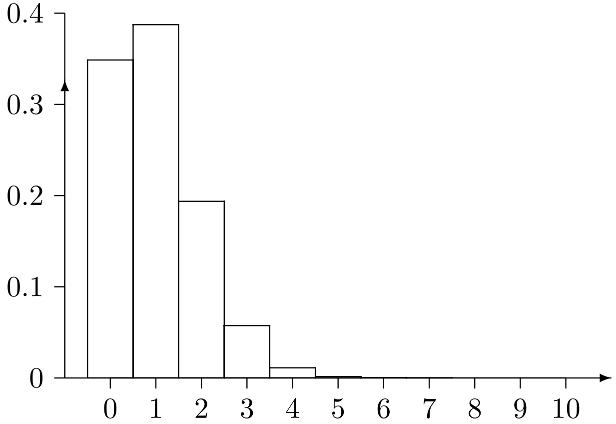
Zum Begiff $Parameter$:

Dies ist eine Größe, die die Wahrscheinlichkeitsverteilung „steuert“.

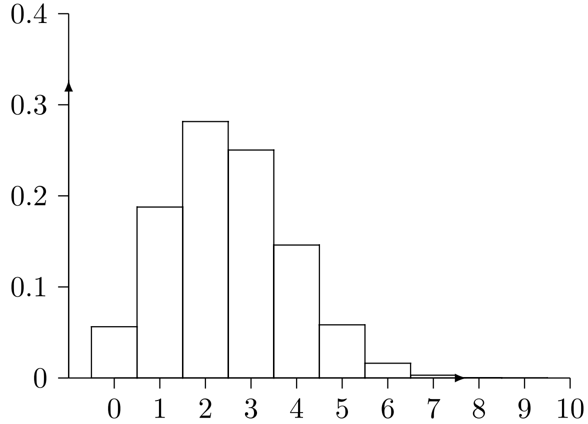
Ist der Parameter bekannt, so kennt man –wenn man die Modellklasse als bekannt voraussetzt– die Wahrscheinlichkeitsverteilung. Später, in der induktiven Statistik, spricht man dann von „parametrischen Modellen“, d.h. der Verteilungstyp wird als bekannt/gegeben („Verteilungsannahme“) vorausgesetzt (z.B. dadurch, dass man durch die Versuchsanordnung mit n unabhängigen Versuchen ein Binomialmodell erzeugt), der konkrete Wert des Parameters bzw. allgemeine Aussagen über seinen Wert sind dann anhand der Stichprobe zu bestimmen.

“Wahrscheinlichkeitshistogramme“ von Binomialverteilungen mit $n = 10$: (graphische Darstellung der Wahrscheinlichkeitsfunktion)

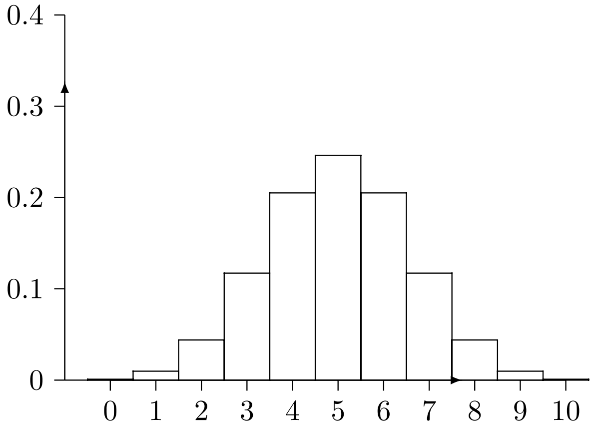
$\pi = 0.1$



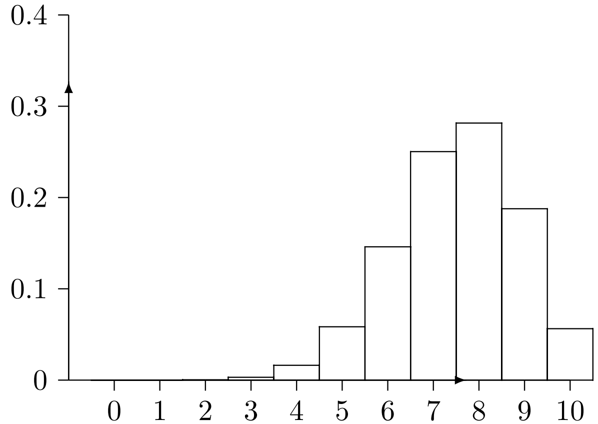
$\pi = 0.25$



$\pi = 0.5$



$\pi = 0.75$



Erwartungswert und Varianz:

- Zur Berechnung von Erwartungswert und Varianz der Binomialverteilung ist folgende Darstellung hilfreich:
$$X = X_1 + \ldots + X_n$$
mit den binären Variablen
$$X_i = \begin{cases} 1 & \text{falls } A \text{ beim } i\text{-ten Versuch eintritt,} \\ 0 & \text{sonst.} \end{cases}$$

- Die X_i sind stochastisch unabhängig mit

$$\begin{aligned} E(X_i) &= 0 \cdot P(X_i = 0) + 1 \cdot P(X_i = 1) = \pi \\ \mathrm{Var}(X_i) &= E(X_i^2) - (E(X_i))^2 = 1 \cdot P(X_i = 1) - \pi^2 = \pi - \pi^2 = \pi(1 - \pi). \end{aligned}$$

- Erwartungswert der Binomialverteilung:

$$E(X) = E(X_1 + \ldots + X_n) = E(X_1) + \ldots + E(X_n) = n\pi$$

Die direkte Berechnung über

$$E(X) = \sum_{i=1}^{n} i \cdot P(\{X = i\}) = \sum_{i=1}^{n} i \binom{n}{i} \pi^i (1-\pi)^{n-i} = \ldots = n\pi$$

ist deutlich komplizierter!

- Varianz der Binomialverteilung:

$$\mathrm{Var}(X) = \mathrm{Var}(X_1 + \ldots + X_n) = \mathrm{Var}(X_1) + \ldots + \mathrm{Var}(X_n) = n\pi(1-\pi)$$

Bsp. 1.76.

Risikobereite Slolalomfahrer stürzen mit Wahrscheinlichkeit 10%, vorsichtigere mit 2%.

a) Schlagen Sie ein Modell für diese Situation vor und diskutieren Sie kurz die zugrunde gelegten Annahmen.

b) Wie groß sind jeweils die Wahrscheinlichkeiten, dass von 20 Fahrern mindestens einer stürzt?

c) Vergleichen Sie die durchschnittlich zu erwartende Anzahl von Stürzen von je 100 Rennläufern!

Exkurs: Zur Problematik der Argumentation mittels „natürlicher Häufigkeiten“ (vgl. v.a. Kap 1.2 und Kap 1.3):
Es wurde wiederholt vorgeschlagen, Wahrscheinlichkeiten anschaulich über „natürliche Häufigkeiten“ zu kommunizieren, also z. B. $P(A) = 0.3753$ darstellen als „von 10000 Personen haben 3753 die Eigenschaft A“.
Man würde demgemäß die Wahrscheinlichkeit $\pi_r = 0.1$ kommunizieren als „von 100 stürzen 10 Rennläufer“.

Diese Darstellung läuft Gefahr, die beträchtliche Variabilität zufälliger Prozesse zu verschleiern. In der Tat ist hier die Wahrscheinlichkeit, dass genau 10 von 100 Läufern stürzen,

$$P(X = 10) = \binom{100}{10} \cdot 0.1^{10} \cdot 0.9^{90} \approx 0.13,$$

also lediglich etwa 13%. „Natürliche Häufigkeiten“ müssen also unbedingt als Durchschnittswerte bzw. Erwartungswerte begriffen und kommuniziert werden.

Eigenschaften der Binomialverteilung:

- Symmetrieeigenschaft (vertausche Rolle von A und $\bar{A}$):

- Summeneigenschaft: Seien $X \sim B(n, \pi)$ und $Y \sim B(m, \pi)$. Sind X und Y unabhängig, so gilt

$$X + Y \sim$$

Tabellierung der Binomialverteilung: Tabelliert ist oft $P(X \leq x)$

$\pi = 0.3$	$n =$11	$n =$12	...
$x \leq 0$	0.0198	0.0138	
1	0.1130	0.0850	...
2	0.3127	0.2528	
$\vdots$	$\vdots$	$\vdots$	

Daraus lassen sich die interessierenden Wahrscheinlichkeiten ablesen:

$$P(X = x) = P(X \leq x) - P(X \leq x - 1), \quad x \in \mathbb{N}_0$$

Zum Beispiel:

$$\begin{aligned} P(X = 2) &= P(X \leq 2) - P(X \leq 1) = \\ &= 0.1997 \\ \Bigg(&= \binom{11}{2} \cdot 0.3^2 \cdot 0.7^2. \Bigg) \end{aligned}$$

Wegen der Symmetrieeigenschaft gibt es meist nur Tabellen für $\pi \leq 0.5$.

Für großes n verwendet man Approximationen durch die Normalverteilung (vgl. Abschnitt **??**).