

2 Zufallsvariablen

Wir gehen hier zunächst davon aus, dass der Ereignisraum Ω eines Zufallsexperimentes höchstens abzählbar unendlich ist.

2.1 Grundlagen

Definition 2.1. Sei Ω der Ereignisraum eines Zufallsexperimentes. Eine Abbildung

$$X : \Omega \rightarrow \mathbb{R}$$

heißt **Zufallsvariable**. Das Bild von X nennen wir die möglichen Realisierungen der Zufallsvariablen. Ist auf Ω eine W.K. definiert, ist also (Ω, P) ein Wahrscheinlichkeitsraum, dann können wir mit Hilfe einer Zufallsvariable eine Abbildung $F : \mathbb{R} \rightarrow [0, 1]$ wie folgt definieren:

$$F(x) := P(\{\omega \in \Omega : X(\omega) \leq x\}).$$

Wir nennen F die zugehörige **Verteilungsfunktion**. Die Verteilungsfunktion ordnet also gewissen Ereignissen eine Wahrscheinlichkeit zu. Statt $P(\{\omega \in \Omega : X(\omega) \leq x\})$ schreibt man meistens einfacher $P(X(\omega) \leq x)$ oder $P(X \leq x)$. Da im Falle endlicher Ereignisräume Ω auch nur endlich viele Realisierungen existieren, können wir eine Abbildung $\text{Bild}(X) \rightarrow [0, 1]$ definieren durch

$$x \mapsto P(X(\omega) = x)$$

und wir schreiben $p_x := P(X(\omega) = x)$. Wenn die Realisierungen x_1, x_2, \dots mit i indiziert werden, schreibt man auch p_i statt p_{x_i} . Eine Zufallsvariable heißt **diskret**, wenn es nur abzählbar viele Realisierungen gibt. Da wir hier davon ausgehen, dass der Ereignisraum Ω höchstens abzählbar viele Elemente hat, sind bei uns Zufallsvariablen zunächst einmal automatisch diskret.

Bemerkung 2.2. 1. Achtung: Die p_x sind nicht die Wahrscheinlichkeiten für Elementarereignisse, weil viele Elementarereignisse durch X auf dasselbe Bild abgebildet werden können.

2. Wir schreiben

$$\begin{aligned} X = x &:= \{\omega : X(\omega) = x\} \\ X < x &:= \{\omega : X(\omega) < x\} \\ x < X \leq y &:= \{\omega : x < X(\omega) \leq y\}. \end{aligned}$$

Proposition 2.3. Sei X eine Zufallsvariable auf einem diskreten Wahrscheinlichkeitsraum. Dann gilt

$$P(x < X \leq y) = F(y) - F(x).$$

Ferner ist

$$F(x) = \sum_{\omega \in \Omega : X(\omega) \leq x} P(\omega) = \sum_{y \leq x} p_y.$$

Beispiel 2.4. In der Vorlesung werden wir die Verteilungsfunktion und die p_x der Zufallsvariable “Augensumme beim Würfeln mit zwei Würfeln” diskutieren.

Satz 2.5. Für eine Verteilungsfunktion F gilt:

(1.) F ist monoton wachsend.

(2.) Für jedes $x^* \in \mathbb{R}$ existieren die Grenzwerte $\lim_{x \nearrow x^*} F(x)$ sowie $\lim_{x \searrow x^*} F(x)$ und es gilt

$$\lim_{x \searrow x^*} F(x) = F(x^*),$$

d.h. F ist rechtsseitig stetig.

(3.) $\lim_{x \searrow x^*} F(x) - \lim_{x \nearrow x^*} F(x) = P(X = x^*)$.

(4.) $\lim_{x \rightarrow -\infty} F(x) = 0$.

(5.) $\lim_{x \rightarrow +\infty} F(x) = 1$.

Bemerkung 2.6. Der Punkt (3.) zeigt, dass wir die p_x unmittelbar aus der Verteilungsfunktion ablesen können und umgekehrt.

Definition 2.7. Ist X eine Zufallsvariable mit zugehöriger Verteilungsfunktion F , so sagen wir, die Z.V. ist F -verteilt, geschrieben $X \sim F$.

Bemerkung 2.8. Beachten Sie, dass die Verteilungsfunktion eine Eigenschaft der Zufallsvariablen ist, **nicht** eines Wahrscheinlichkeitsraumes.

Wir werden später sehen, dass man jeder Zufallsvariablen gewisse Zahlen zuordnen kann, die wichtige Eigenschaften der Z.V. beschreiben. Ein erstes Beispiel sind die Quantile:

Definition 2.9. Sei X eine Zufallsvariable und $p \in (0, 1)$. Ferner sei F die zugehörige Verteilungsfunktion. Eine Zahl Q_p mit

$$\lim_{x \nearrow Q_p} F(x) \leq p \leq F(Q_p)$$

heißt p -**Quantil** der Zufallsvariablen. Im Fall $p = 1/2$ spricht man vom **Median**.

Actung: Die Zahlen Q_p sind nicht eindeutig bestimmt. Wenn F eine Umkehrfunktion hat, gilt $Q_p = F^{-1}(p)$.

2.2 Binomialverteilung, Hypergeometrische Verteilung, Poissonverteilung

Die einfachste Verteilung ist die **Gleichverteilung**, bei der $P(X = x_i) = 1/N$ gilt, wenn N die Anzahl möglicher Realisierungen x_1, \dots, x_N von X bezeichnet. Die zugehörige Verteilungsfunktion ist eine Treppenfunktion. Klassisches Beispiel: Münzwurf oder würfeln.

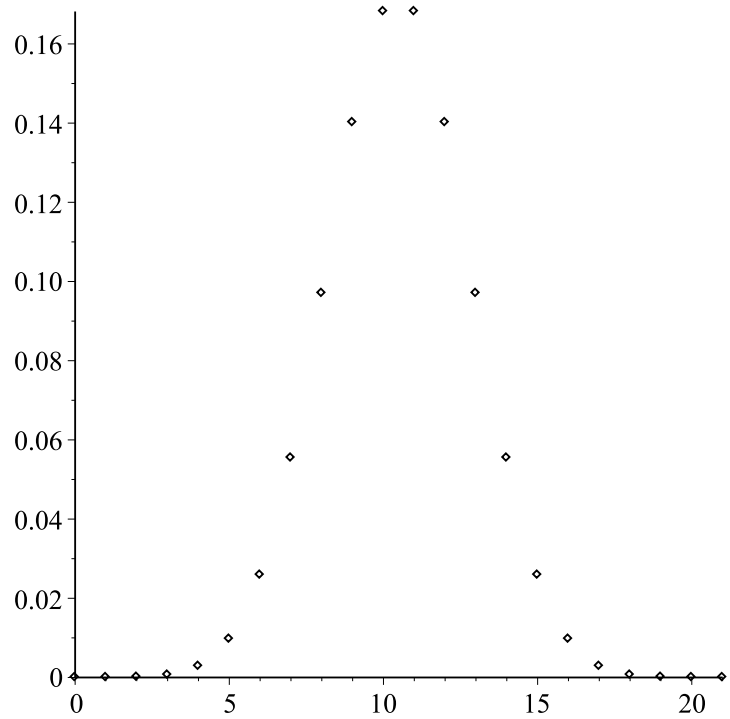
Definition 2.10. Ein **Bernoulli-Experiment** ist ein Zufallsexperiment, bei dem es nur zwei mögliche Ausgänge A und B gibt. Das Ereignis A trete mit W.K. p und das Ereignis B dann mit W.K. $1 - p = q$ ein. Damit wird $\{A, B\}$ zu einem Wahrscheinlichkeitsraum. Wir haben kein Laplace-Experiment, es sei denn, $p = q = 1/2$.

Wenn wir ein Bernoulliexperiment n mal wiederholen, ist der Ereignisraum $\Omega = \{A, B\}^n$. Die Elementarereignisse heißen Bernoulliketten. Eine Kette mit k Ereignissen A und $n - k$ Ereignissen B tritt mit W.K. $p^k q^{n-k}$ ein. So wird auf Ω ein Wahrscheinlichkeitsmaß definiert. Wir definieren nun eine Zufallsvariable X auf diesem Wahrscheinlichkeitsraum wie folgt: Einer Bernoullikette mit k Einträgen A und $n - k$ Einträgen B wird die Zahl k (manchmal auch “Anzahl Erfolge” genannt) zugeordnet. Offenbar gilt

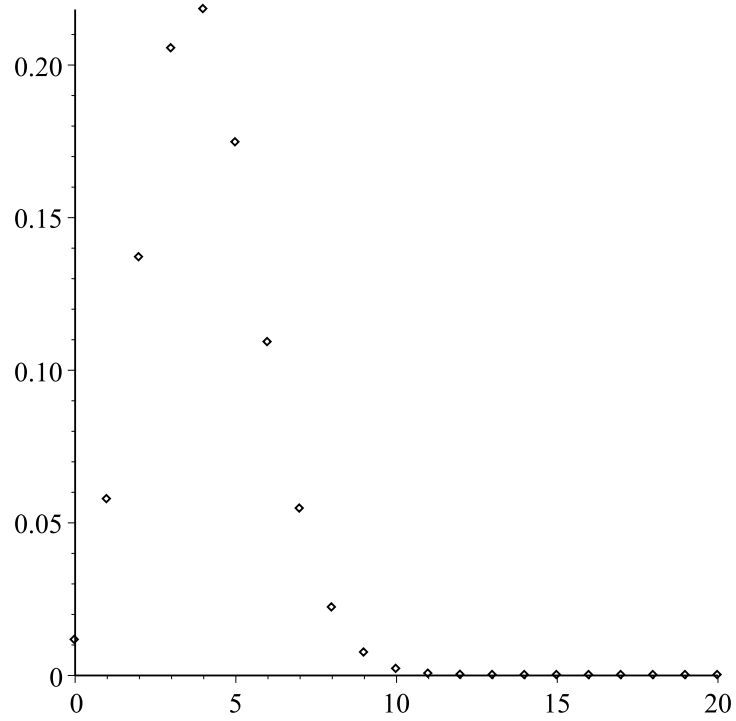
$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$

Dies liefert eine Wahrscheinlichkeitsverteilung, die sogenannte **Binomialverteilung** $B(n, p)$. Sie ist komplett durch Angabe von n und p bestimmt. Wir nennen eine Zufallsvariable binomialverteilt mit Parametern n und p , wenn ihre Wahrscheinlichkeitsverteilung $B(n, p)$ ist. Die folgenden Bilder zeigen die Wahrscheinlichkeitsverteilungen für einige Werte von p und n :

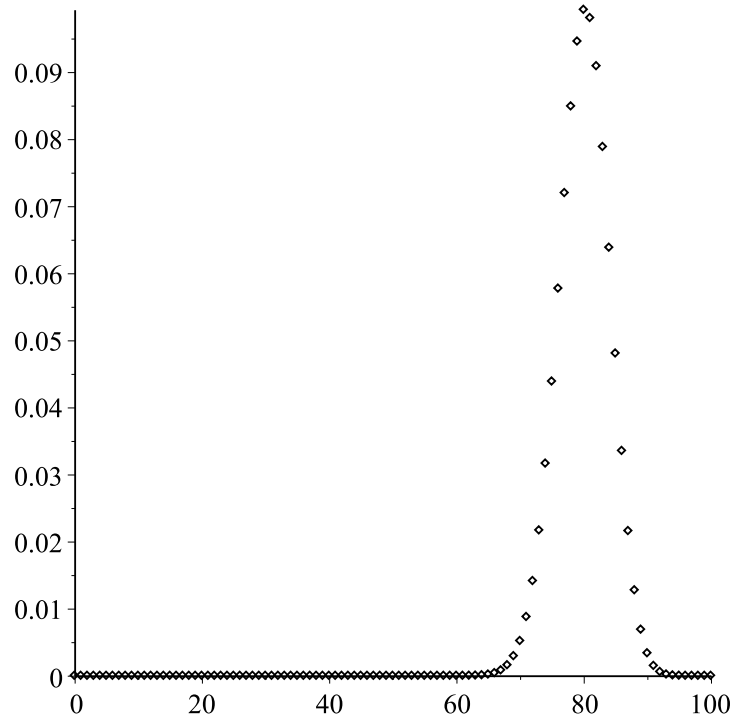
$B(21, 0.5)$:



$B(20, 0.2)$:



$B(100, 0.8)$:



Kommen wir nun zur **hypergeometrischen Verteilung** $H(n, M, N)$: Wir haben hier eine Urne mit N Kugeln, M davon seien rot. Wir ziehen ohne Zurücklegen n Kugeln. Diesem Zufallsexperiment können wir ein Wahrscheinlichkeitsmaß zuordnen sowie eine Z.V. X definieren, nämlich die Anzahl k roter Kugeln. Es gilt

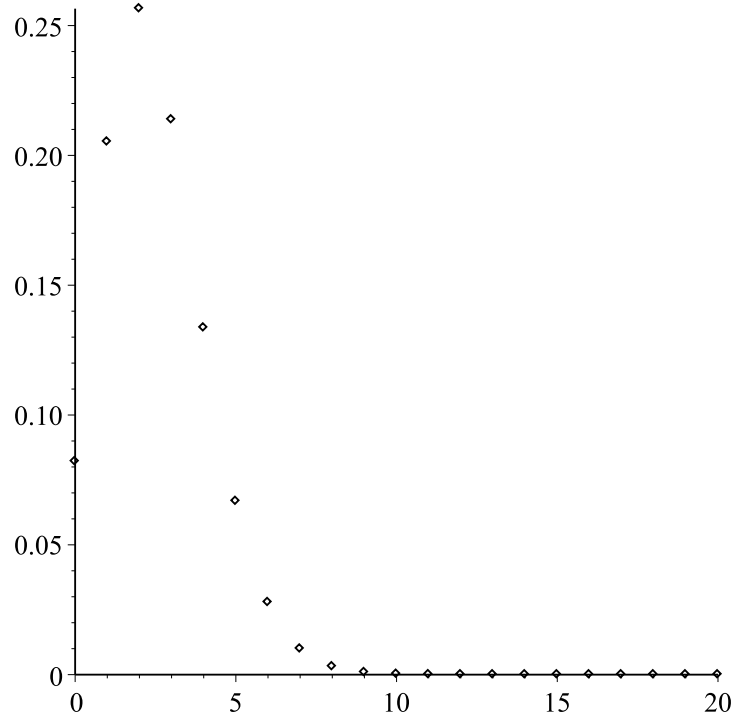
$$P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}.$$

Eine Zufallsvariable mit dieser Wahrscheinlichkeitsverteilung heißt hypergeometrisch verteilt.

Abschließend betrachten wir die **Poisson-Verteilung**. Der Ereignisraum ist hier $\{0, 1, \dots\}$ die Menge der nicht-negativen ganzen Zahlen. Die Poissonverteilung modelliert recht gut die Anzahl seltener Ereignisse, die in einem fest gewählten Zeitraum auftreten (Tore pro Fußballspiel). Wir nennen eine Zufallsvariable X Poisson-verteilt $P(\lambda)$ mit Parameter λ , wenn

$$P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}.$$

Hier ist ein Bild für $\lambda = 2.5$, was etwa der Anzahl geschossener Tore in einem Bundesligaspiel entspricht:



Interessant ist, dass die Poissonverteilung als Grenzverteilung der Binomialverteilung interpretiert werden kann: Gilt $X \sim B(n, p)$ mit großem n und kleinem p , so können wir die Approximation

$$P(X = k) \approx \frac{\lambda^k}{k!} e^{-\lambda}.$$

nutzen. Faustregel: $n \geq 50$ und $p \leq 0.1$.

2.3 Lagemaße

Jeder Zufallsvariablen kann man gewisse Zahlen zuordnen kann, die wichtige Eigenschaften der Z.V. beschreiben. Ein Beispiel haben wir bereits gesehen, die Quantile (Definition 2.9). In diesem Abschnitt folgen Erwartungswert und Varianz.

Definition 2.11. Sei X eine Zufallsvariable mit Realisierungen x_1, x_2, \dots . Der

Erwartungswert $E(X)$ von X ist definiert als

$$E(X) := \sum_{i=1}^{\infty} x_i \cdot P(X = x_i)$$

sofern diese Summe existiert (was im Fall *endlicher* Wahrscheinlichkeitsräume immer der Fall ist).

Bemerkung 2.12. Es gilt $E(X) = \sum_{\omega \in \Omega} X(\omega)P(\omega)$.

Beispiel 2.13. Wenn wir mit zwei Würfeln werfen und als Zufallsvariable jedem Ausgang den Betrag der Differenz der Augenzahlen zuordnen, so ist der Erwartungswert $35/18$, siehe Vorlesung.

Definition 2.14. Sei X eine Zufallsvariable mit Realisierungen x_1, x_2, \dots und Erwartungswert μ . Wir definieren die Varianz

$$V(X) := E((X - \mu)^2)$$

sofern dieser Erwartungswert existiert.

Bemerkung 2.15. Es gilt

$$V(X) = \sum_{i=1}^{\infty} (x_i - \mu)^2 \cdot P(X = x_i)$$

sowie

$$V(X) = E(X^2) - \mu^2.$$

Ferner wird die Wurzel aus der Varianz auch Standardabweichung σ genannt. Entsprechend schreibt man für die Varianz manchmal σ^2 .

Man kann die Abweichung einer Z.V. vom Mittelwert mit Hilfe der Varianz abschätzen:

Satz 2.16 (Tschebyscheff'sche Ungleichung).

$$P(|X - E(X)| \geq \epsilon) \leq \frac{1}{\epsilon^2} V(X).$$

Diese Abschätzung gilt für beliebige Zufallsvariablen und ist wegen ihrer Allgemeinheit nicht sehr stark.

Wir haben drei wichtige Wahrscheinlichkeitsverteilungen kennengelernt: Binomial, Poisson, hypergeometrisch. In der folgenden Tabelle fassen wir Erwartungswert und Varianz dieser Verteilungen zusammen:

| | $E(X)$ | $V(X)$ |
|--------------|-----------------|---|
| $B(n, p)$ | np | $np(1-p)$ |
| $H(n, M, N)$ | $n \frac{M}{N}$ | $n \frac{M}{N} (1 - \frac{M}{N}) \frac{N-n}{N-1}$ |
| $P(\lambda)$ | λ | λ |

Bemerkung 2.17. Wenn X eine Zufallsvariable ist, dann ist auch $g(X)$ für eine Funktion $g : \mathbb{R} \rightarrow \mathbb{R}$ eine Zufallsvariable. Man kann aber nicht unmittelbar Erwartungswert, Varianz und Wahrscheinlichkeitsverteilung aus g ablesen. Es gilt

$$E(g(X)) = \sum_{i=0}^{\infty} g(x_i)P(X = x_i).$$

Satz 2.18. Sei X eine Zufallsvariable mit Erwartungswert $E(X) = \mu$ und Varianz $V(X) = \sigma^2$. Dann gilt

$$\begin{aligned} E(aX + b) &= a \cdot E(X) + b \\ V(aX + b) &= a^2 \cdot V(X). \end{aligned}$$

Insbesondere ist $\frac{X - \mu}{\sigma}$ eine Zufallsvariable mit Erwartungswert 0 und Varianz 1 (sofern $\sigma \neq 0$).

2.4 Gemeinsame Verteilungen und Lagemaße von Zufallsvariablen

Satz 2.19. Seien X und Y Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, P) mit Erwartungswerten $E(X)$ und $E(Y)$. Dann gilt

$$E(aX + Y) = aE(X) + E(Y).$$

Ein entsprechender Satz für die Varianz gilt zunächst einmal nicht.

Definition 2.20. Zwei Zufallsvariablen X und Y heißen **unabhängig**, wenn für alle reellen Zahlen x, y gilt:

$$P(X = x \text{ und } Y = y) = P(X = x) \cdot P(Y = y).$$

Mit anderen Worten: Die beiden Ereignisse

$$\begin{aligned} \{\omega : X(\omega) = x\} \\ \{\omega : Y(\omega) = y\} \end{aligned}$$

sind unabhängig. Allgemein nennen wir Z.V. X_1, \dots, X_n auf (Ω, P) unabhängig, wenn für alle $x_1, \dots, x_n \in \mathbb{R}$ gilt

$$P(X_i = x_i \text{ für } i = 1, \dots, n) = \prod_{i=1}^n P(X = x_i).$$

Bemerkung 2.21. Beachten Sie, dass X und Y auf dem selben Wahrscheinlichkeitsraum definiert sind!

Bemerkung 2.22. Unabhängigkeit von mehr als zwei Zufallsvariablen ist eine stärkere Bedingung als paarweise Unabhängigkeit!

Für unabhängige Zufallsvariablen gilt:

Satz 2.23. Seien X und Y unabhängige Zufallsvariablen auf einem W.R. (Ω, P) . Dann gilt

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

sofern beide Erwartungswerte existieren, ebenso für das Produkt mehrerer unabhängiger Zufallsvariablen

Im Fall unabhängiger Z.V. kann man auch etwas über die Summe der Varianz sagen:

Satz 2.24. Seien X und Y unabhängige Zufallsvariablen auf einem W.R. (Ω, P) . Dann gilt

$$V(X + Y) = V(X) + V(Y)$$

sofern $V(X), V(Y)$ existieren.

Bemerkung 2.25. Beide Sätze gelten auch für mehr als zwei unabhängige Zufallsvariablen.

Definition 2.26. Die Kovarianz zweier Zufallsvariablen X, Y ist definiert als

$$C(X, Y) = E(X \cdot Y) - E(X) \cdot E(Y)$$

sofern die entsprechenden Erwartungswerte existieren.

Die Bedeutung der Kovarianz wird im folgenden Satz deutlich:

Satz 2.27. Seien X und Y zwei Zufallsvariablen auf dem W.R. (Ω, P) . Dann gilt

$$V(X + Y) = V(X) + V(Y) + 2C(X, Y).$$

Eigenschaften der Kovarianz:

Proposition 2.28. $X, Y, X_1, \dots, Y_1, \dots$ seien Z.V. auf W.R. (Ω, P) . Dann gilt

(1.) $C(X, X) = V(X), \quad C(X, Y) = C(Y, X).$

(2.) $C(X, Y) = E((X - E(X)) \cdot (Y - E(Y))).$

(3.) $C(X + a, Y + b) = C(X, Y).$

(4.) Wenn X und Y unabhängig sind, dann gilt $C(X, Y) = 0.$

(5.) $V(X_1 + \dots + X_n) = \sum_{i=1}^n V(X_i) + 2 \cdot \sum_{1 \leq i < j \leq n} C(X_i, X_j) .$

(6.) $C(\sum_i a_i X_i, \sum_j b_j Y_j) = \sum_{i,j} a_i b_j C(X_i, Y_j).$

$$(7.) C(X, Y)^2 \leq V(X) \cdot V(Y).$$

Bemerkung 2.29. Die letzte Eigenschaft zeigt, dass wir die Kovarianz normieren können. Wir nennen den Quotienten

$$-1 \leq \frac{C(X, Y)}{\sqrt{V(X) \cdot V(Y)}} \leq 1$$

die **Korrelation** zwischen X und Y . Liegt die Korrelation nahe bei ± 1 , so besteht zwischen X und Y eine "große" Abhängigkeit. Unabhängige Z.V. sind unkorreliert (d.h. Korrelation = 0), aber nicht umgekehrt: Man kann sich beispielsweise überlegen, dass Summe und Differenz beim zweimaligen Würfeln unkorrelierte Z.V. sind, die aber nicht unabhängig sind.

2.5 Produkträume

Was passiert, wenn wir Z.V. X_i haben, die auf verschiedenen Wahrscheinlichkeitsräumen (Ω_i, P_i) definiert sind. Dazu definieren wir den Produktraum $\Omega = \Omega_1 \times \dots \times \Omega_n$ und definieren auf Ω das Produktmaß

$$P(\omega_1, \dots, \omega_m) = \prod_{i=1}^m P_i(\omega_i).$$

Zumindest im abzählbar unendlichen Fall ist das kein Problem und wir können so jeder Teilmenge von Ω eine Wahrscheinlichkeit zuordnen. Die X_i kann man als unabhängige Zufallsvariablen auf diesem Produktraum auffassen und wir können damit arbeiten so wie oben für den Fall dass die verschiedenen Z.V. auf **einem** W.R. definiert sind. Klassischer Fall: Zufallsexperiment wird n -mal wiederholt. Dann bezeichnet (Ω_i, P_i) den Wahrscheinlichkeitsraum, der zur i -ten Durchführung des Zufallsexperimentes gehört.

Es gilt das für die Statistik wichtige schwache Gesetz der großen Zahlen:

Satz 2.30. *Es seien X_1, \dots, X_n unabhängige Zufallsvariablen auf einem Wahrscheinlichkeitsraum (Ω, P) , die alle dieselbe Verteilungsfunktion haben (i.i.d.: independently identical distribution) mit Erwartungswert μ und Varianz σ^2 . Dann ist auch $\tilde{X}_n = \frac{1}{n}(X_1 + \dots + X_n)$ eine Zufallsvariable auf Ω mit*

$$\lim_{n \rightarrow \infty} P(|\tilde{X}_n - \mu| < \epsilon) = 1$$

für jedes $\epsilon > 0$.

Bemerkung 2.31. Der klassische Fall ist der, dass ein Zufallsexperiment mit Z.V. X mehrfach (n mal) durchgeführt wird. Formal haben wir dann aber n unabhängige Z.V. X_1, \dots, X_n , die jeweils den Ausgang des i -ten Experimentes beschreiben. Formal sind das dann unabhängige Z.V. auf dem Produktraum. Ferner kann man die Voraussetzungen etwas abschwächen.