Erwartungswert einer transformierten ZV $g(X)$,
wobei $g: \mathbb{R} \longrightarrow \mathbb{R}$ eine reellwertige Funktion ist.

Bsp.: Berechnung der Varianz (s. nächster Abschnitt) erfordert $E[X^2]$ mit $g(X) = X^2$

Satz 3.1:

Sei $Y = g(X)$ eine Funktion der ZV X. Dann gilt

(1) für eine diskrete ZV mit Wahrscheinlichkeitsfkt. $p(x)$: $\quad E[g(X)] = \sum_i g(x_i)\, p(x_i)$

(2) für eine stetige ZV mit Wahrscheinlichkeitsdichte $f(x)$: $\quad E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx$

Bemerkung: Dieser Satz ist hilfreich, wenn man nur an dem Lagemaß einer transformierten ZV interessiert ist und nicht an ihrer Verteilung.

Beispiel: Die ZV X sei die Zeit in Stunden, die benötigt wird, um einen ausgefallenen Rechner in einem Rechnernetz wieder funktionsfähig zu machen. Die Dichtefunktion von X sei gegeben durch

$$f(x) = \begin{cases} 1 & , \; 0 < x < 1 \\ 0 & , \; \text{sonst} \end{cases}$$

Die Kosten $g(X) = X^3$ (in Hundert €) für einen Rechnerausfall sind abhängig von der Ausfallzeit X.

Wie hoch sind die durchschnittlich zu erwartenden Kosten?

$$E[g(X)] = \int_{-\infty}^{\infty} g(x)\, f(x)\, dx = \int_0^1 x^3\, dx = \left[\tfrac{1}{4}x^4\right]_0^1 = \frac{1}{4} \text{ von Hundert €}$$

Eigenschaften des Erwartungswerts:

konstante ZV

$$\boxed{\begin{array}{lll} (1) & E[b] & = b \\ (2) & E[aX+b] & = a\cdot E[X]+b \\ (3) & E[X_1+X_2+\ldots+X_n] & = \sum_{i=1}^{n} E[X_i] \end{array}}$$

zu (2): $E[aX+b] \overset{\text{diskreter Fall}}{=} \sum_i (ax_i+b)\cdot p(x_i) \overset{\text{Linearität der Summe}}{=} a \underbrace{\sum_i x_i\, p(x_i)}_{E[X]} + b \cdot \underbrace{\sum_i p(x_i)}_{=1}$

$= a\,E[X] + b$

Beispiel: Auf 20 verschiedene Prozessoren, die mit gleicher Wahrscheinlichkeit ausgewählt werden, sollen $k = 10$ Stapelaufträge verteilt werden. Berechnen Sie den Erwartungswert für die Anzahl verschiedener Prozessoren in der Auswahl.

Die Anzahl Y verschiedener Prozessoren lässt sich beschreiben durch

Gegenereignis zu keinmal ausgewählt

$$Y = \sum_{i=1}^{20} X_i \qquad \text{mit} \qquad X_i = \begin{cases} 1, & \text{falls Prozessor } i \text{ mind. einmal ausgewählt wurde} \\ 0, & \text{sonst} \end{cases}$$

"Indikatorfunktion" des Ereignisses "Prozessor i ausgewählt"

Ges.: $E[Y] = E\left[\sum_{i=1}^{20} X_i\right] \overset{(3)}{=} \sum_{i=1}^{20} E[X_i] = 20\cdot\left(1-\left(\frac{19}{20}\right)^{10}\right) \approx 8$

$$E[X_i] = 1\cdot p(1) + 0\cdot p(0) = 1\cdot\left(1-\left(\frac{19}{20}\right)^{10}\right) \approx 0{,}4$$

3.4 Varianz

Die Varianz bzw. die Standardabweichung ist die wichtigste Maßzahl für das Streuungsverhalten einer Verteilung.

Definition 3.9: Varianz / Standardabweichung

Die Varianz ist die mittlere quadratische Abweichung einer ZV X von ihrem Erwartungswert $E[X]=\mu$ und ist definiert durch

$$\sigma^2 = Var[X] = E\left[(X-\mu)^2\right]$$

Die Standardabweichung ist die Wurzel aus der Varianz

$$\sigma = \sqrt{Var[X]}$$

und hat die gleiche Einheit wie X und μ.

Berechnung der Varianz:

- diskret: $Var[X] = \sum_i (x_i-\mu)^2 \cdot p(x_i)$
- stetig: $Var[X] = \int_{-\infty}^{\infty} (x-\mu)^2 \cdot f(x)\, dx$

- **Satz 3.2: Verschiebungssatz für die Varianz**

$$\boxed{Var[X] = E[X^2] - (E[X])^2}$$

Beweis: 6. Übung

Eigenschaften der Varianz:

(1) $Var[b] = 0$

(2) $Var[aX+b] = a^2 \cdot Var[X]$

(3) $Var[X_1 + \ldots + X_n] =$ s. Abschnitt 3.5

3.5 Kovarianz und Varianz einer Summe von ZVen

Definition 3.10: Kovarianz

Die Kovarianz $Cov[X,Y]$ zweier Zufallsvariablen X und Y wird definiert durch

$$Cov[X,Y] = E\big[(X-E[X])(Y-E[Y])\big]$$

Bemerkung: Die Kovarianz beschreibt die Abhängigkeit zweier ZV X und Y. Je größer die Korrelation zw. X und Y, desto größer (betragsmäßig) ist die Kovarianz.

Eigenschaften der Kovarianz:

(1) $Cov[X,Y] = Cov[Y,X]$

(2) $Cov[X,X] = Var[X]$

(3) $Cov[aX,Y] = a \cdot Cov[X,Y]$

Satz 3.3: Verschiebungssatz für die Kovarianz

$$Cov[X,Y] = E[X \cdot Y] - E[X] \cdot E[Y]$$

Beweis: 6. Übung

Satz 3.4: Varianz von Summen

$$Var[X_1 + X_2 + \ldots + X_n] = \sum_{i=1}^{n} Var[X_i] + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} Cov[X_i, X_j]$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} Cov[X_i, X_j]$$

$n=2$: $$Var[X_1 + X_2] = Var[X_1] + Var[X_2] + \underbrace{Cov[X_1,X_2] + Cov[X_2,X_1]}_{2 \cdot Cov[X_1,X_2]}$$

Eine deutliche Vereinfachung ergibt sich für stochastisch unabh. ZVen

$$X_1, X_2 \text{ unabhängig} \quad \underset{\not\Leftarrow}{\Longrightarrow} \quad Cov[X_1, X_2] = 0$$

Dann gilt:

$$\boxed{Var[X_1 + \ldots + X_n] = \sum_{i=1}^{n} Var[X_i]}$$

Beispiel: Ein Kommunikationssystem besteht aus 10 Komponenten, die voneinander unabhängig mit der Wahrscheinlichkeit $p = 0.1$ funktionieren.

Gesucht ist die Varianz für die ZV Y = "Anzahl der funktionsfähigen Komponenten".

$$Y = \sum_{i=1}^{10} X_i \quad \text{mit} \quad X_i = \begin{cases} 1 & \text{, falls Komponente } i \text{ funkt.} \\ 0 & \text{, falls Komponente } i \text{ nicht funkt.} \end{cases}$$

Ges.: $$Var[Y] = Var\left[\sum_{i=1}^{10} X_i\right] \overset{\text{Unabh. der Komponenten}}{=} \sum_{i=1}^{10} Var[X_i] = 10 \cdot 0{,}09 = 0{,}9$$

$$Var[X_i] \underset{\text{Verschiebungssatz}}{=} E[X_i^2] - (E[X_i])^2 = 0{,}1 - 0{,}1^2 = 0{,}09$$

$$E[X_i] = 1 \cdot p(1) + 0 \cdot p(0) = 0{,}1$$

$$E[X_i^2] = 1^2 \cdot p(1) + 0^2\, p(0) = 0{,}1$$

3.6 Quantile

Definition 3.11: p-Quantil

Sei X eine reellwertige ZV mit Verteilungsfkt. $F(x)$ und sei $0<p<1$. Dann ist das p-Quantil der kleinste Wert $x_p \in \mathbb{R}$, für den gilt:

$$F(x_p) \geq p \, .$$

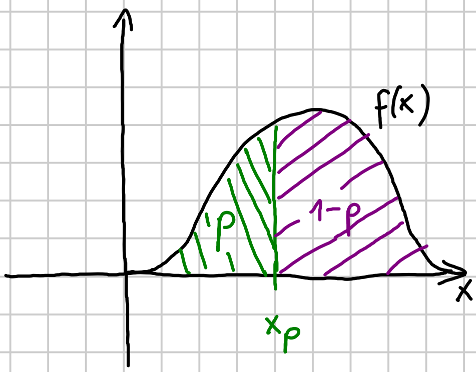
Spezielle Quantile:

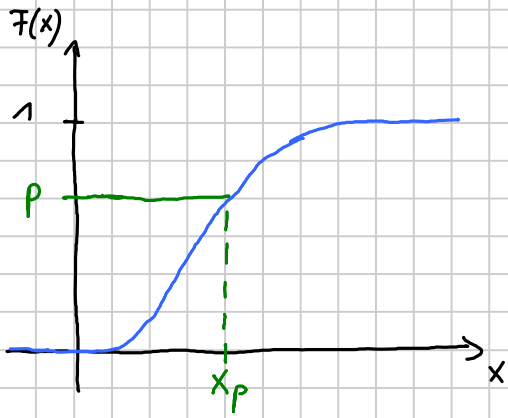
$p = 0{,}25$: 1. Quartil

$p = 0{,}5$: 2. Quartil oder Median

$p = 0{,}75$: 3. Quartil

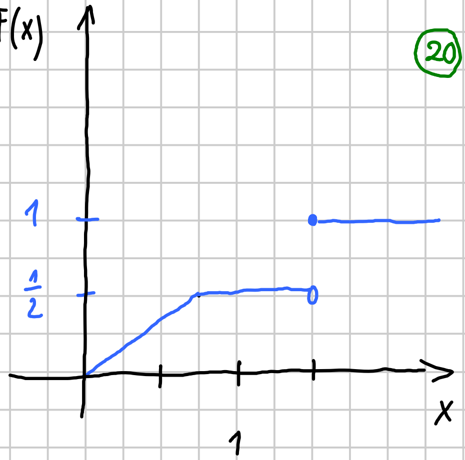
Bemerkung: Bei einer stetigen ZV teilt das p-Quantil x_p die Fläche unterhalb der Dichtefunktion f in die Anteile p (links von x_p) und $1-p$ (rechts von x_p) auf. Falls die Verteilungsfkt. F streng monoton und damit umkehrbar ist, dann gilt: $x_p = F^{-1}(p)$





Beispiel:

Geg.: $F(x) = \begin{cases} 0 & , x < 0 \\ \frac{2}{3}x & , 0 \leq x < 3/4 \\ \frac{1}{2} & , \frac{3}{4} \leq x < \frac{3}{2} \\ 1 & , \frac{3}{2} \leq x \end{cases}$



Bestimmen Sie das 1., 2. und 3. Quartil.

$x_{0,25}$: $F(x_{0,25}) = \boxed{0,25 = \frac{2}{3} \cdot x_{0,25}} \Rightarrow x_{0,25} = \frac{1}{4} \cdot \frac{3}{2} = \frac{3}{8} = 0,375$

$x_{0,5}$: $F(x_{0,5}) = 0,5 \xRightarrow{\text{Def. 3.11}}$ Kleinster x-Wert für den $F(x) \geq 0,5$

$\Rightarrow x_{0,5} = \frac{3}{4} = 0,75$

$x_{0,75}$: Kleinster x-Wert, für den $F(x) \geq 0,75 \Rightarrow x_{0,75} = \frac{3}{2}$

3.7 Chebyshev-Ungleichung und schwaches Gesetz der großen Zahlen

Satz 3.5: Chebyshev-Ungleichung

Sei X eine ZV mit Erwartungswert μ und Varianz σ^2.

Dann gilt für ein beliebige reelle Zahl $k > 0$:

$$P(|X - \mu| \geq k) \leq \frac{\sigma^2}{k^2}.$$

Die Chebyshev-Ungleichung hat eher eine theoretische als eine praktische Bedeutung, da sie für Abschätzungen in Beweisen verwendet wird, z.B. bei folgendem wichtigen Satz.

Satz 3.6: Schwaches Gesetz der großen Zahlen

Sei $X_1, X_2, \ldots$ eine Folge von unabhängigen, identisch verteilten ZVen mit $E[X_i] = \mu$ und $Var[X_i] = \sigma^2$.

Dann gilt für ein bel. kleines $\varepsilon > 0$:

$$P\left(\left|\underbrace{\frac{X_1 + \ldots + X_n}{n}}_{\bar{X}} - \mu\right| > \varepsilon\right) \xrightarrow[n \to \infty]{} 0$$

In Worten: Der Mittelwert $\bar{X}$ konvergiert stochastisch gegen den EW μ.
(die Wahrscheinlichkeit geht gegen 0)

Beweis:

1. Schritt: Zu zeigen: $E[X_i] = \mu$ ist auch EW von $\frac{1}{n}\sum_{i=1}^{n} X_i = \bar{X}$

$$E\left[\frac{1}{n}\cdot\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\, E\left[\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n}\cdot n\mu = \mu$$

2. Schritt: Verwendung der Chebyshev-Ungleichung

$$P\left(|\bar{X} - \mu| > \varepsilon\right) \leq \frac{Var[\bar{X}]}{\varepsilon^2} = \frac{\frac{\sigma^2}{n}}{\varepsilon^2} = \frac{1}{n}\cdot \underbrace{\frac{\sigma^2}{\varepsilon^2}}_{\text{konstant}} \xrightarrow[n \to \infty]{} 0$$

$$Var[\bar{X}] = Var\left[\frac{1}{n}\cdot\sum_{i=1}^{n} X_i\right] = \frac{1}{n^2}\, Var\left[\sum_{i=1}^{n} X_i\right] \overset{X_i \text{ unabh.}}{=} \frac{1}{n^2}\sum_{i=1}^{n} Var[X_i]$$

$$= \frac{1}{n^2}\cdot n\sigma^2 = \frac{\sigma^2}{n}$$