Regressionsgerade

$$\widehat{y} = \widehat{m} \cdot x + \widehat{b}$$

Bestimmtheitsmaß

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \widehat{y}_i)^2}{\sum_{i=1}^n (y_i - \overline{y})^2} = \frac{\sum_{i=1}^n (\widehat{y}_i - \overline{y})^2}{\sum_{i=1}^n (y_i - \overline{y})^2} = \frac{s_{\widehat{y}}^2}{s_y^2} \quad \text{mit } \widehat{y}_i = \widehat{m} \cdot x_i + \widehat{b}, i = 1, \ldots, n\,.$$

R^2: Anteil an der Gesamtvarianz s_y^2, der durch die Regression erklärt werden kann.
Es gilt: $R^2 = r_{xy}^2$.
Stochastische Unabhängigkeit von zwei Ereignissen A und B: $P(A \cap B) = P(A) \cdot P(B)$
Formel der totalen Wahrscheinlichkeit: Sei $A \subset \Omega$, $\Omega = B_1 \cup \cdots \cup B_k$, $B_j \cap B_i = \emptyset$ für $j \neq i$. Dann gilt

$$\mathsf{P}(A) \;=\; \mathsf{P}(A|B_1) \cdot \mathsf{P}(B_1) \;+\; \cdots \;+ \mathsf{P}(A|B_k) \cdot \mathsf{P}(B_k)$$

Bayessche Formel: Für Ereignisse $A, B \subset \Omega$ mit $P(A) > 0$ und $P(B) > 0$ gilt

$$\mathsf{P}(B|A) \;=\; \frac{\mathsf{P}(A|B) \cdot \mathsf{P}(B)}{\mathsf{P}(A)}\;, \quad \text{wobei } \mathsf{P}(A|B) := \frac{\mathsf{P}(A \cap B)}{\mathsf{P}(B)} \text{ bedingte Wahrscheinlichkeit von } A \text{ gegeben } B$$

Binomialverteilung mit den Parametern n und p

$$\mathsf{P}(X = k) := \binom{n}{k} p^k (1-p)^{n-k}\;, \quad \text{wobei } \binom{n}{k} \;=\; \frac{n!}{k!\,(n-k)!}\;, \quad k \in \{0, 1, 2, \cdots, n\},$$

Poisson-Verteilung mit dem Parameter $\lambda > 0$

$$\mathsf{P}(X = k) := e^{-\lambda} \cdot \frac{\lambda^k}{k!}\;, \quad k \in \{0, 1, 2, \cdots\} =: \mathbb{N},$$

Normalverteilung Sei $X \sim \mathcal{N}(\mu, \sigma^2)$. Dann gilt ($\Phi$ Verteilungsfunktion der Standardnormalverteilung):

$$P(a < X \leq b) = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

Falls die folgenden Reihen und (uneigentlichen) Integrale existieren, definiert man: **Erwartungswert** und **Varianz** einer **diskreten** Zufallsvariable X:

$$\mathsf{E}X := \sum_{i=1}^{\infty} x_i \cdot \mathsf{P}(X = x_i) \text{ und } \mathsf{Var}X := \sum_{i=1}^{\infty} (x_i - \mathsf{E}X)^2 \cdot \mathsf{P}(X = x_i) = \sum_{i=1}^{\infty} x_i^2 \cdot \mathsf{P}(X = x_i) - (\mathsf{E}X)^2$$

Erwartungswert und **Varianz** einer **stetigen** Zufallsvariable X:

$$\mathsf{E}X := \int_{-\infty}^{\infty} x \cdot f(x)dx \text{ und } \mathsf{Var}X := \int_{-\infty}^{\infty} (x - \mathsf{E}X)^2 \cdot f(x)dx = \int_{-\infty}^{\infty} x^2 \cdot f(x)dx - (\mathsf{E}X)^2$$

Wahrscheinlichkeitsdichte: Funktion $f : \mathbb{R} \to [0, \infty)$ mit
i) $\forall x \in \mathbb{R} : f(x) \geq 0$
ii) $\int_{-\infty}^{\infty} f(x)dx = 1$
Für jede Dichte f gilt: $\int_{-\infty}^{x} f(t)dt = F(x)$, wobei F die Verteilungsfunktion zur Dichte f ist, mit $F(x) = \mathsf{P}(X \leq x)$.
$(1-\alpha)$-Konfidenzintervall für den Erwartungswert μ einer normalverteilten ZV bei bekannter Varianz σ^2:

$$\left[\overline{x} \;-\; u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\;,\; \overline{x} \;+\; u_{1-\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right].$$

Hierbei ist $u_{1-\frac{\alpha}{2}}$ das $(1 - \frac{\alpha}{2})$ - Quantil der Standardnormalverteilung.
$(1-\alpha)$-Konfidenzintervall für den Erwartungswert μ einer normalverteilten ZV bei unbekannter Varianz:

$$\left[\overline{x} \;-\; t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\;,\; \overline{x} \;+\; t_{n-1;1-\frac{\alpha}{2}} \frac{s}{\sqrt{n}}\right].$$

Hierbei ist $t_{n-1;1-\frac{\alpha}{2}}$ das $(1 - \frac{\alpha}{2})$ - Quantil der t-Verteilung mit $n - 1$ Freiheitsgraden.