# Formelsammlung Statistik

**Lageparameter und Streumaße:**

**Einfaches arithmetisches Mittel:** $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \left[ oder \quad \overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \right]$

**Gewogenes arithmetisches Mittel:**

$$\mu = \frac{\sum_{i=1}^{k}(x_i \cdot n_i)}{\sum_{i=1}^{k} n_i} = \frac{1}{n}\sum_{i=1}^{k}(x_i \cdot n_i) = \sum_{i=1}^{k}\left(x_i \cdot \frac{n_i}{n}\right)$$

**Arithmetisches Mittel bei klassierten Merkmalsausprägungen:**

$$\mu = \frac{1}{n}\sum_{i=1}^{k}\left[(x_i)_m \cdot n_i\right] = \sum_{i=1}^{k}\left[(x_i)_m \cdot \frac{n_i}{n}\right] \quad mit\ (x_i)_m\ als\ Klassenmitte\ der\ Klasse\ i$$

**Median (Zentralwert):** $$\overline{x_M} = \begin{cases} x_{\frac{n+1}{2}} & \text{für n ungerade} \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right) & \text{für n gerade} \end{cases}$$

**Quantile:** $$\overline{x_p} = \begin{cases} x_{[n \cdot p]+1} & \text{für } (\mathrm{n} \cdot \mathrm{p}) \text{ nicht ganzzahlig} \\ \frac{1}{2}\left(x_{n \cdot p} + x_{n \cdot p+1}\right) & \text{für } (\mathrm{n} \cdot \mathrm{p}) \text{ ganzzahlig} \end{cases}$$

**Varianz:** $\sigma^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \mu)^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i)^2 - \mu^2$

**Varianz bei Häufigkeitsverteilung:** $\sigma^2 = \frac{1}{n}\sum_{i=1}^{k}(x_i - \mu)^2 \cdot n_i$

**Standardabweichung:** $\sigma = \sqrt{\sigma^2}$

***Modus:*** *Häufigster Wert einer Verteilung.*

**Gini-Koeffizient:**

*Berechung der Flächen mit relativen kumulierten Häufigkeiten:*

$$Gini \quad = \quad \frac{\text{Konzentrationsfläche K}}{\text{maximale Konzentrationsfläche } \mathrm{K}_{\max}}$$

Anmerkung: Als *„Konzentrationsfläche"* *K* bezeichnet man die Fläche, zwischen der Gleichverteilungsdiagonalen und der LORENZ-Kurve: $0 \le K \le 1/2$ .

$$Wegen \quad K_{\max} = \frac{1}{2} - \frac{1}{2n} \quad lautet\ der\ normierte\ Gini-Koeffizient:$$

$$norm.\,Gini \quad = \quad K \cdot \frac{2n}{n-1}$$

***Der Wertebereich des Gini-Koeffizienten liegt zwischen 0 (= Gleichverteilung) und 1 (= vollständige Konzentration auf einen Merkmalsträger).***

**Preisindizes**

**nach Laspeyres:** $$L_P \quad = \quad \frac{\sum p_{1i} \cdot q_{0i}}{\sum p_{0i} \cdot q_{0i}} \quad = \quad \sum \frac{p_{1i}}{p_{0i}} \cdot \frac{p_{0i} \cdot q_{0i}}{\sum p_{0i} \cdot q_{0i}} \qquad mit\ i \in N$$

**nach Paasche:** $$P_P \quad = \quad \frac{\sum p_{1i} \cdot q_{1i}}{\sum p_{0i} \cdot q_{1i}} \quad = \quad \sum \frac{p_{1i}}{p_{0i}} \cdot \frac{p_{0i} \cdot q_{1i}}{\sum p_{0i} \cdot q_{1i}} \qquad mit\ i \in N$$

**Lineare Regression und Korrelation**

**<u>Ansatz:</u>** $$y = b_0 + b_1 x$$

$$\mu_x = \frac{1}{n}\sum_{i=1}^{n} x_i \qquad \mu_y = \frac{1}{n}\sum_{i=1}^{n} y_i$$

$$b_1 = \frac{\sigma_{xy}}{\sigma_{x^2}} = \frac{\sum_{i=1}^{n}(x_i y_i) - n\cdot\mu_x\cdot\mu_y}{\sum_{i=1}^{n}\left(x_i^{\,2}\right) - n\cdot\mu_x^{\,2}} = \frac{\frac{1}{n}\sum_{i=1}^{n}(x_i y_i) - \mu_x\cdot\mu_y}{\frac{1}{n}\sum_{i=1}^{n}\left(x_i^{\,2}\right) - \mu_x^{\,2}}$$

$$b_0 = \mu_y - b_1\cdot\mu_x$$

**Korrelation nach Brevais-Pearson:**

$$r = \frac{\sum_{i=1}^{n}(x_i-\mu_x)\cdot(y_i-\mu_y)}{\sqrt{\sum_{i=1}^{n}(x_i-\mu_x)^2}\cdot\sqrt{\sum_{i=1}^{n}(y_i-\mu_y)^2}} = \frac{\sum_{i=1}^{n}(x_i y_i) - n\cdot\mu_x\cdot\mu_y}{\sqrt{\sum_{i=1}^{n}\left(x_i^{\,2}\right) - n\cdot\mu_x^{\,2}}\cdot\sqrt{\sum_{i=1}^{n}\left(y_i^{\,2}\right) - n\cdot\mu_y^{\,2}}}$$

**Rangkorrelation nach Spearman**

$$r_S = 1 - \frac{6\sum_{i=1}^{n}(d_i)^2}{n\left(n^2-1\right)}$$

**Stochastik**

**Erwartungswert einer diskreten Zufallsvariablen:**

$$\mu = E(X) = \sum_{i=1}^{n} x_i \cdot P(X = x_i) \quad mit \quad X = \{x_1, x_2, ..., x_n\}$$

**Stochastische Unabhängigkeit:**

$$A, B \text{ sind stochastisch unabhängig} \Leftrightarrow P(A \cap B) = P(A) \cdot P(B)$$

**Bedingte Wahrscheinlichkeit:**

$$P_B(A) = \frac{P(A \cap B)}{P(B)}$$ W'keit für A unter der Bedingung B (B ist bereits eingetreten)

**Totale Wahrscheinlichkeit:**

$$P(B) = \sum_{i=1}^{k} \left[ P_{A_i}(B) \cdot P(A_i) \right]$$

**Satz von Bayes:**

$$P_B(A_j) = \frac{P_{A_j}(B) \cdot P(A_j)}{\sum_{i=1}^{k} \left[ P_{A_i}(B) \cdot P(A_i) \right]}$$

**Dichtefunktionen:**

Eine Funktion f(x) heißt Wahrscheinlichkeitsdichte (Dichtefunktion) genau dann, wenn gilt:

$$(1) \quad f(x) \geq 0 \quad \forall x \in \Re \qquad und \qquad (2) \quad \int_{-\infty}^{\infty} f(x)\,dx = 1$$

Ist X eine stetige Zufallsgröße mit der Dichtefunktion f(x), so heißt die reelle Zahl

(i) $\mu(X) = \int_{-\infty}^{\infty} x \cdot f(x)\,dx$ der Erwartungswert der Zufallsgröße X.

(ii) $V(X) = \int_{-\infty}^{\infty} (x - \mu)^2 \cdot f(x)\,dx$ die Varianz der Zufallsgröße X.

**Diskrete Wahrscheinlichkeitsverteilungen**

**Hypergeometrische Verteilung:**

$$H(X=k) = \frac{\binom{M}{k}\binom{N-M}{n-k}}{\binom{N}{n}} \quad mit \quad 0 \le k \le n \quad und \quad M \le N \quad und \quad n \le N \quad und \quad k \le M$$

**Binomialverteilung:**

$$B(X=k) = \binom{n}{k} p^k \cdot (1-p)^{n-k} \quad mit \quad k \le n$$

Erwartungswert der Zufallsvariablen: $\mu = n \cdot p$

Varianz der Zufallsvariablen: $\sigma^2 = n \cdot p \cdot (1-p) = \mu \cdot (1-p)$

**Poissonverteilung:**

$$P(X=k) = \frac{\mu^k}{k!} e^{-\mu} \quad mit \quad \mu = n \cdot p$$

**Stetige Wahrscheinlichkeitsverteilung - (Standard)Normalverteilung:**

**Wahrscheinlichkeitsdichte / Gauß-Glockenfunktion:**

$$\varphi_{\mu;\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\cdot\left(\frac{x-\mu}{\sigma}\right)^2} \xrightarrow[und \quad \sigma=1]{mit \quad \mu=0} \varphi_{0;1}(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

=> $P(X \le k) = \Phi_{\mu;\sigma}\left(\frac{k-\mu}{\sigma}\right) \quad und \quad x = \frac{k-\mu}{\sigma}$

=> $\Phi(-x) = 1-\Phi(x) \quad und \quad \Phi(x)-\Phi(-x) = 2\Phi(x)-1$

## Gaußsche Summenfunktion

| x | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|---|
| **0,0** | 0,500 | 0,504 | 0,508 | 0,512 | 0,516 | 0,520 | 0,524 | 0,528 | 0,532 | 0,536 |
| **0,1** | 0,540 | 0,544 | 0,548 | 0,552 | 0,556 | 0,560 | 0,564 | 0,567 | 0,571 | 0,575 |
| **0,2** | 0,579 | 0,583 | 0,587 | 0,591 | 0,595 | 0,599 | 0,603 | 0,606 | 0,610 | 0,614 |
| **0,3** | 0,618 | 0,622 | 0,626 | 0,629 | 0,633 | 0,637 | 0,641 | 0,644 | 0,648 | 0,652 |
| **0,4** | 0,655 | 0,659 | 0,663 | 0,666 | 0,670 | 0,674 | 0,677 | 0,681 | 0,684 | 0,688 |
| **0,5** | 0,691 | 0,695 | 0,698 | 0,702 | 0,705 | 0,709 | 0,712 | 0,716 | 0,719 | 0,722 |
| **0,6** | 0,726 | 0,729 | 0,732 | 0,736 | 0,739 | 0,742 | 0,745 | 0,749 | 0,752 | 0,755 |
| **0,7** | 0,758 | 0,761 | 0,764 | 0,767 | 0,770 | 0,773 | 0,776 | 0,779 | 0,782 | 0,785 |
| **0,8** | 0,788 | 0,791 | 0,794 | 0,797 | 0,800 | 0,802 | 0,805 | 0,808 | 0,811 | 0,813 |
| **0,9** | 0,816 | 0,819 | 0,821 | 0,824 | 0,826 | 0,829 | 0,831 | 0,834 | 0,836 | 0,839 |
| **1,0** | 0,841 | 0,844 | 0,846 | 0,848 | 0,851 | 0,853 | 0,855 | 0,858 | 0,860 | 0,862 |
| **1,1** | 0,864 | 0,867 | 0,869 | 0,871 | 0,873 | 0,875 | 0,877 | 0,879 | 0,881 | 0,883 |
| **1,2** | 0,885 | 0,887 | 0,889 | 0,891 | 0,893 | 0,894 | 0,896 | 0,898 | 0,900 | 0,901 |
| **1,3** | 0,903 | 0,905 | 0,907 | 0,908 | 0,910 | 0,911 | 0,913 | 0,915 | 0,916 | 0,918 |
| **1,4** | 0,919 | 0,921 | 0,922 | 0,924 | 0,925 | 0,926 | 0,928 | 0,929 | 0,931 | 0,932 |
| **1,5** | 0,933 | 0,934 | 0,936 | 0,937 | 0,938 | 0,939 | 0,941 | 0,942 | 0,943 | 0,944 |
| **1,6** | 0,945 | 0,946 | 0,947 | 0,948 | 0,949 | 0,951 | 0,952 | 0,953 | 0,954 | 0,954 |
| **1,7** | 0,955 | 0,956 | 0,957 | 0,958 | 0,959 | 0,960 | 0,961 | 0,962 | 0,962 | 0,963 |
| **1,8** | 0,964 | 0,965 | 0,966 | 0,966 | 0,967 | 0,968 | 0,969 | 0,969 | 0,970 | 0,971 |
| **1,9** | 0,971 | 0,972 | 0,973 | 0,973 | 0,974 | 0,974 | 0,975 | 0,976 | 0,976 | 0,977 |
| **2,0** | 0,977 | 0,978 | 0,978 | 0,979 | 0,979 | 0,980 | 0,980 | 0,981 | 0,981 | 0,982 |
| **2,1** | 0,982 | 0,983 | 0,983 | 0,983 | 0,984 | 0,984 | 0,985 | 0,985 | 0,985 | 0,986 |
| **2,2** | 0,986 | 0,986 | 0,987 | 0,987 | 0,987 | 0,988 | 0,988 | 0,988 | 0,989 | 0,989 |
| **2,3** | 0,989 | 0,990 | 0,990 | 0,990 | 0,990 | 0,991 | 0,991 | 0,991 | 0,991 | 0,992 |
| **2,4** | 0,992 | 0,992 | 0,992 | 0,992 | 0,993 | 0,993 | 0,993 | 0,993 | 0,993 | 0,994 |
| **2,5** | 0,994 | 0,994 | 0,994 | 0,994 | 0,994 | 0,995 | 0,995 | 0,995 | 0,995 | 0,995 |
| **2,6** | 0,995 | 0,995 | 0,996 | 0,996 | 0,996 | 0,996 | 0,996 | 0,996 | 0,996 | 0,996 |
| **2,7** | 0,997 | 0,997 | 0,997 | 0,997 | 0,997 | 0,997 | 0,997 | 0,997 | 0,997 | 0,997 |
| **2,8** | 0,997 | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 |
| **2,9** | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 | 0,998 | 0,999 | 0,999 | 0,999 |