

1. Mathematik, Folgen und Reihen, Potenzreihen
2. Mathematik, Differential- und Integralrechnung
3. Binomial- und Hypergeometrische Verteilung, Approximationen
4. Normalverteilung, Berechnen von Wahrscheinlichkeiten
5. Chi-Quadrat-Verteilung, Erwartungswert, Varianz
6. Nullhypothese, Alternativhypothese, Tests, p-Werte
7. Lineare Algebra, Matrizenrechnung
8. t-Test und Varianzanalyse
Produkt von Matrizen
9. Eigenwerte, Eigenvektoren

3a. Bsp. $X \sim Bi(n, p)$, $p = 0.1, n = 25$

exakt (Binomial):

$$\begin{aligned} P(X \geq 5) &= \sum_{k=5}^{25} \binom{n}{k} p^k (1-p)^{n-k} \\ &= 1 - P(X \leq 4) = 1 - \text{CDF('Binomial', 4, p, n)} \\ &= 0.097994 \end{aligned}$$

approximativ (ZGWS, normal)

$$\begin{aligned} P(X > 4) &= 1 - P(X \leq 4) \\ &= 1 - P\left(\frac{X - np}{\sqrt{np(1-p)}} \leq \frac{4 - np}{\sqrt{np(1-p)}}\right) \\ &\approx 1 - \Phi\left(\frac{1.5}{5 \cdot 0.3}\right) = 1 - \Phi(1) = \Phi(-1) \\ &= \text{CDF('Normal', -1, 0, 1)} = 0.15866 \end{aligned}$$

$$\begin{aligned} P(X \geq 5) &= 1 - P(X < 5) \\ &\approx 1 - P\left(\frac{X - np}{\sqrt{np(1-p)}} < \frac{5 - np}{\sqrt{np(1-p)}}\right) \\ &= 1 - \Phi\left(\frac{5 - 2.5}{5 \cdot 0.3}\right) \\ &= 1 - \Phi\left(\frac{5}{3}\right) = \Phi\left(-\frac{5}{3}\right) = 0.04779 \end{aligned}$$

approximativ (ZGWS, normal, Stetigkeitskorrektur):

$$\begin{aligned} P(X > 5) &\approx 1 - P\left(\frac{X - np - 0.5}{\sqrt{np(1-p)}} \leq \frac{5 - np - 0.5}{\sqrt{np(1-p)}}\right) \\ &= 1 - \Phi\left(\frac{4.5 - 2.5}{5 \cdot 0.3}\right) = 1 - \Phi\left(\frac{4}{3}\right) \\ &= \Phi\left(-\frac{4}{3}\right) = 0.09121 \end{aligned}$$

Die Approximation der Binomial- durch eine Normalverteilung ist hier nicht so gut (vor allem ist $p = 0.1$ klein).

3b. Hypergeometrische und Binomial Verteilung

Seien

$$f(k|Bi(m, p)) = \binom{m}{k} p^k (1-p)^{m-k} \quad \text{Binomialwkt.}$$

und

$$f(k|H_{N,n,m}) = \frac{\binom{n}{k} \cdot \binom{N-n}{m-k}}{\binom{N}{m}} \quad \text{Hypergeom.wkt.}$$

Satz: Es gilt Für $N \rightarrow \infty$, $n \rightarrow \infty$, $\frac{n}{N} \rightarrow p$ gilt:

$$f(k|H_{N,n,m}) \rightarrow \binom{m}{k} p^k (1-p)^{m-k} = f(k|Bi(m, p))$$

Beweis: Es gilt

$$\begin{aligned}
f(k|H_{N,n,m}) &= \frac{\binom{n}{k} \cdot \binom{N-n}{m-k}}{\binom{N}{m}} \\
&= \frac{n! \cdot (N-n)! \cdot m!(N-m)!}{k!(n-k)! \cdot (N-n-m+k)!(m-k)! \cdot N!} \\
&= \binom{m}{k} \frac{(N-n)!}{(N-n-m+k)!(n-k)!} \frac{n!}{N!} \frac{(N-m)!}{N!} \\
&= \binom{m}{k} (N-n-m+k+1) \cdots (N-n) \cdot \\
&\quad \frac{(n-k+1) \cdots n}{(N-m+1) \cdots N} \\
&= \binom{m}{k} \frac{(N-n-m+k+1) \cdots (N-n)}{N^{m-k}} \cdot \\
&\quad \frac{(n-k+1) \cdots n}{N^k} \frac{N^m}{(N-m+1) \cdots N} \\
&= \binom{m}{k} \left(1-p - \frac{m-k-1}{N}\right) \left(1-p - \frac{m-k-2}{N}\right) \cdots \left(1-p - \frac{m-1}{N}\right) \\
&\quad \frac{\left(p - \frac{k-1}{N}\right) \cdots p}{\left(1 - \frac{m-1}{N}\right) \left(1 - \frac{m-2}{N}\right) \cdots \left(1 - \frac{1}{N}\right) \cdot 1} \\
&\rightarrow \binom{m}{k} (1-p)^{m-k} p^k = f(k|Bi(m,p))
\end{aligned}$$

5a. χ^2 -Verteilung

Def.: Seien $X_i \sim \mathcal{N}(0, 1), i = 1, \dots, n$, und unabhängig. Dann ist

$$Y = \sum_{i=1}^n X_i^2 \sim \chi_n^2$$

χ^2 -verteilt mit n Freiheitsgraden.

Erwartungswert und Varianz lassen sich leicht ausrechnen:

$$\begin{aligned} \mathbf{E}(Y) &= n\mathbf{E}(X)^2 = \text{var}(X) + (\mathbf{E}(X))^2 = n(1 + 0) = n \\ \text{var}(Y) &= n \cdot \text{var}(X^2) = n(\mathbf{E}X^4 - (\mathbf{E}(X^2))^2) \\ &= n(3 - 1^2) = 2n \end{aligned}$$

Anmerkung: Die Dichte von Y ist gegeben durch ($y > 0$)

$$f_{\chi_n^2}(y) = \frac{1}{2^{n/2}\Gamma(\frac{n}{2})} e^{-\frac{y}{2}} y^{\frac{n}{2}-1}.$$

5b. Sei $X \sim \text{Exp}(\lambda)$. Es gilt:

$$\mathbf{E}X = \lambda \quad \text{und} \quad \text{var}X = \lambda^2.$$

Beweis:

$$\begin{aligned} \mathbf{E}X &= \int_0^{\infty} x \cdot \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx = \\ &= x(-e^{-\frac{x}{\lambda}})|_0^{\infty} - \int_0^{\infty} (-e^{-\frac{x}{\lambda}}) dx \\ &= \lambda. \end{aligned}$$

$$\begin{aligned} \mathbf{E}X^2 &= \int_0^{\infty} x^2 \frac{1}{\lambda} e^{-\frac{x}{\lambda}} dx = \\ &= x^2(-e^{-\frac{x}{\lambda}})|_0^{\infty} - \int_0^{\infty} 2x(-e^{-\frac{x}{\lambda}}) dx \\ &= 2\lambda^2 \\ \text{var}X &= 2\lambda^2 - \lambda^2 = \lambda^2. \end{aligned}$$

6. Im zweiseitigen Fall ist $H_0 : \mu = \mu_0$. $H_A : \mu \neq \mu_0$,
d.h. der Parameterraum ist $\Theta = \mathbb{R}$.

Im einseitigen Fall gibt es folgende Varianten

$$H_0 : \mu \leq \mu_0 \text{ und } H_A : \mu > \mu_0,$$

d.h. $\Theta = \mathbb{R}$ (V) oder

$$H_0 : \mu = \mu_0 \text{ und } H_A : \mu > \mu_0,$$

d.h. $\Theta = \{\mu : \mu \geq \mu_0\}$ (ÜA 12a)

Beide Fälle werden gleich behandelt.

Wenn t die Realisierung von T ist und $t < 0$ dann
wird H_0 ohnehin nicht abgelehnt, der p-Wert ist größer
als 0.5.

Der Fall

$$H_0 : \mu \geq \mu_0 \text{ und } H_A : \mu < \mu_0 \text{ oder}$$

$$H_0 : \mu = \mu_0 \text{ und } H_A : \mu < \mu_0$$

ist analog.

7. Matrizenmultiplikation ist assoziativ, nicht kommutativ.

Sei $\mathbf{X} = (X_1, \dots, X_n)$ der Beobachtungsvektor. Sei \mathbf{I} die Einheitsmatrix und $\mathbf{1}$ die Matrix, die nur aus Einsen besteht. Es gilt (bitte nachrechnen)

$$(n-1)s^2 = \sum_{i=1}^n (X_i - \bar{X})^2 = \mathbf{X}^T A \mathbf{X},$$

wobei

$$A = \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T$$

Weiterhin gilt: A ist symmetrisch und (bitte nachrechnen)

$$A^2 = A$$

Bem.: Matrizen mit dieser Eigenschaft heißen idempotent.

Seien die $X_i \sim \mathcal{N}$ und die Matrix A idempotent.

Dann sind Quadratische Formen $\mathbf{X}^T A \mathbf{X} \sim \chi_{fg}^2$ mit $fg = rg(A)$ Freiheitsgraden.

8. Betrachten wir die F -Statistik aus der einfaktoriellen Varianzanalyse

$$F = \frac{MSB}{MSE} = \frac{N-k}{k-1} \frac{SSB}{SSW} = \frac{N-k}{k-1} \frac{\sum_{i=1}^k n_i (\bar{Y}_i - \bar{Y})^2}{\sum_{i=1}^k \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2}$$

und setzen die Anzahl der Behandlungen auf $k = 2$.

Dann wird mit $\bar{Y} = \frac{1}{N}(n_1\bar{Y}_1 + n_2\bar{Y}_2)$, $N = n_1 + n_2$

$$F = \frac{N-2}{2-1} \left(\frac{n_1(\bar{Y}_1 - \frac{1}{N}(n_1\bar{Y}_1 + n_2\bar{Y}_2))^2}{(n_1-1)s_1^2 + (n_2-1)s_2^2} + \frac{n_2(\bar{Y}_2 - \frac{1}{N}(n_1\bar{Y}_1 + n_2\bar{Y}_2))^2}{(n_1-1)s_1^2 + (n_2-1)s_2^2} \right)$$

Die Zähler in den rechten Brüchen sind zusammen gleich

$$\begin{aligned} &= n_1 \left(\left(1 - \frac{n_1}{N}\right) \bar{Y}_1 - \frac{n_2}{N} \bar{Y}_2 \right)^2 + n_2 \left(-\frac{n_1}{N} \bar{Y}_1 + \left(1 - \frac{n_2}{N}\right) \bar{Y}_2 \right)^2 \\ &= \frac{n_1 n_2^2}{N^2} (\bar{Y}_1 - \bar{Y}_2)^2 + \frac{n_1^2 n_2}{N^2} (\bar{Y}_1 - \bar{Y}_2)^2 \\ &= \frac{n_1 n_2}{N} (\bar{Y}_1 - \bar{Y}_2)^2 \end{aligned}$$

Also:

$$\begin{aligned} F &= \frac{(N-2)n_1n_2}{N} \frac{(\bar{Y}_1 - \bar{Y}_2)^2}{(n_1-1)s_1^2 + (n_2-1)s_2^2} \\ &= \frac{1}{\frac{1}{n_1} + \frac{1}{n_2}} \frac{(\bar{Y}_1 - \bar{Y}_2)^2}{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{N-2}} \end{aligned}$$

Vergleichen Sie das mit der t -Statistik (gemeinsame Varianzen)

$$T = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{1}{n} + \frac{1}{m}} \sqrt{\frac{(n-1)S_1^2 + (m-1)S_2^2}{n+m-2}}}$$

und das ist dasselbe wie das Quadrat von T .

Bem.: Wir wissen

$$t = \frac{\mathcal{N}(0, 1)}{\sqrt{\frac{1}{n+m-2} \chi_{n+m-2}^2}}$$

(so ist die t -Verteilung definiert)

und es ist

$$t^2 = \frac{\mathcal{N}(0, 1)^2}{\frac{1}{n+m-2} \chi_{n+m-2}^2} = \frac{\frac{1}{1} \chi_1^2}{\frac{1}{n+m-2} \chi_{n+m-2}^2} = F_{1, n+m-2}$$

(so ist die F -Verteilung definiert).