Der folgende Aufschrieb enthält die fehlenden Beweise aus der Vorlesung.

Wiederholung: Transformationsformel für Dichten

Zur Herleitung der χ^2 und t-Verteilung verwenden wir mehrmals die Transformationsformel. Deswegen wird diese hier nochmal kurz wiederholt.

Theorem 0.1 *Sei $X = (X_1, \dots, X_n)^T$ eine Zufallsvariable mit Werten in $\mathbb{R}^n$ und Dichtefunktion $f_X : \mathbb{R}^d \to \mathbb{R}_+$ wobei $\mathbb{R}_+ := \{x \in \mathbb{R} \,|\, x \geq 0\}$ und $\phi : \mathbb{R}^d \to \mathbb{R}^d$ stetig differenzierbar mit stetig differenzierbarer Umkehrabbildung ϕ^{-1}. Dann gilt fuer die Dichte f_Y von $Y = \phi(X)$,*

$$f_Y(y) \;=\; f_X(\phi^{-1}(y))\,|\det J_{\phi^{-1}}(y)|.$$

Proof: Für $A \subset \mathbb{R}^d$ gilt:

$$\int_A f_Y(y)\,dy = \mathrm{P}(Y \in A) = \mathrm{P}(X \in \phi^{-1}(A)) = \int_{\phi^{-1}(A)} f_X(x)\,dx.$$

Es gilt $X = \phi^{-1}(Y)$ und da ϕ als auch ϕ^{-1} stetig differenzierbar sind, können wir die Transformationsformel aus der mehrdimensionalen Integralrechnung anwenden

$$\int_{\phi^{-1}(A)} f_X(x)\,dx = \int_A f_X(\phi^{-1}(y))\,|\det J_{\phi^{-1}}(y)|\,dy$$

Da diese Beziehung für jede beliebige Teilmenge A gilt, erhalten wir

$$f_Y(y) = f_X(\phi^{-1}(y))\,|\det J_{\phi^{-1}}(y)|.$$

□

Wir wenden die Transformationsformel für Dichten nun an, um die Verteilung einer affinen Transformation von Gauss-Variablen zu bestimmen. Dieses Resultat benötigen wir später.

Theorem 0.2 *Sei $A \in \mathbb{R}^{m \times n}$ mit $m \leq n$ und der Rang von A sei m. Ferner sei $X = (X_1, \dots, X_n)^T$ eine Zufallsvariable mit Werten in $\mathbb{R}^n$ und $X \sim \mathcal{N}(\mu, C)$, wobei $\mu \in \mathbb{R}^d$ der Mittelwert ist und $C \in \mathbb{R}^{d \times d}$ die Kovarianzmatrix. Die Verteilung der affinen Transformation $Y = AX + b$ mit $b \in \mathbb{R}^m$ ist gegeben durch*

$$Y \;\sim\; \mathcal{N}(A\mu + b\,, ACA^T).$$

Proof: Wir erweitern die Matrix $A \in \mathbb{R}^{m \times n}$ durch Hinzufügen von Zeilen zu einer Matrix $T \in \mathbb{R}^{n \times n}$ mit vollem Rang n, d.h.

$$T = \begin{pmatrix} A \\ B \end{pmatrix},$$

wobei $B \in \mathbb{R}^{(n-m) \times n}$ Wir wählen die Zeilenvektoren $b_j \in B$, $j = 1, \dots, n-m$, so, daß gilt

$$\langle a_i, Cb_j \rangle = 0, \qquad \forall\, i = 1, \dots, m, \quad j = 1, \dots, n-m.$$

In Matrixnotation kann man das kurz schreiben als: $ACB^T = 0$. Dies ist möglich, da wir immer eine Teilmenge von m linear unabhängigen Vektoren (die Zeilenvektoren von A) in $\mathbb{R}^n$ zu einer Basis (d.h. n linear unabhängigen Vektoren in $\mathbb{R}^n$) ergänzen können (Basisergänzungssatz aus der MfI2). Hier wählen wir die hinzugefügten Zeilenvektoren aus dem Orthogonalraum vom $\mathrm{span}\{a_1, \dots, a_m\}$ bezüglich des inneren Produckt $\langle u, v \rangle := u^T C v$ (beachte, daß C positiv definit ist).

Weiterhin ergänzen wir den Vektor $b \in \mathbb{R}^m$ zu einem Vektor $c \in \mathbb{R}^n$ indem wir ihn mit Nullen auffüllen. Wir betrachten $Z := \phi(X) = TX + c$. Man beachte, daß die ersten m Komponenten von Z gleich $Y = AX + b$ sind. Da die Matrix T vollen Rang hat, ist T invertierbar und wir erhalten die Umkehrabbildung: $X = \phi^{-1}(Z) = T^{-1}(Z - c)$ und $J_{\phi^{-1}}(z) = T^{-1}$ für alle $z \in \mathbb{R}^n$. Da ϕ als auch ϕ^{-1} stetig differenzierbar sind, gilt mit

$$f_Z(z) = f_X(\phi^{-1}(z))\,|\det J_{\phi^{-1}}(z)| = f_X(\phi^{-1}(z))\,|\det T^{-1}|$$

Wir haben in der MfI2 folgende Regeln für Determinanten hergeleitet:

$$\det A^{-1} = (\det A)^{-1}, \qquad \text{und} \qquad \det A = \det(A^T), \qquad \text{und} \qquad \det(AB) = \det(BA) = \det B \det A.$$

Daraus leiten wir ab

$$|\det T|\sqrt{\det C} = \sqrt{\det T \det T^T}\sqrt{\det C} = \sqrt{\det T \det C \det T^T} = \sqrt{\det(TCT^T)}.$$

Damit erhalten wir

$$\begin{aligned}
f_X(T^{-1}(b-z))\frac{1}{|\det T|} &= e^{-\frac{1}{2}\left\langle T^{-1}(z-c)-\mu, C^{-1}(T^{-1}(z-c)-\mu)\right\rangle} \frac{1}{|\det T|\sqrt{\det C}} \frac{1}{(2\pi)^{\frac{n}{2}}} \\
&= e^{-\frac{1}{2}\left\langle T^{-1}(z-c-T\mu), C^{-1}T^{-1}(z-c-T\mu)\right\rangle} \frac{1}{\sqrt{\det(TCT^T)}} \frac{1}{(2\pi)^{\frac{n}{2}}} \\
&= e^{-\frac{1}{2}\left\langle (z-c-T\mu), (T^{-1})^T C^{-1} T^{-1}(z-c-T\mu)\right\rangle} \frac{1}{\sqrt{\det(TCT^T)}} \frac{1}{(2\pi)^{\frac{n}{2}}} \\
&= e^{-\frac{1}{2}\left\langle (z-c-T\mu), (TCT^T)^{-1}(z-c-T\mu)\right\rangle} \frac{1}{\sqrt{\det(TCT^T)}} \frac{1}{(2\pi)^{\frac{n}{2}}},
\end{aligned}$$

wobei wir im letzten Schritt die Beziehung $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$ für die Inverse eines Matrixprodukts ausgenützt haben. Damit hat $Z = TX + c$ eine Normalverteilung mit Mittelwert $T\mu + c$ und Kovarianzmatrix TCT^T d.h. $Z \sim \mathcal{N}(T\mu + c,\, TCT^T)$. Die neue gesamte Kovarianzmatrix ist gegeben als

$$TCT^T = \begin{pmatrix} A \\ B \end{pmatrix} C \begin{pmatrix} A^T & B^T \end{pmatrix} = \begin{pmatrix} ACA^T & ACB^T \\ BCA^T & BCB^T \end{pmatrix} = \begin{pmatrix} ACA^T & 0 \\ 0 & BCB^T \end{pmatrix},$$

wobei wir im letzten Schritt die spezielle Wahl der Zeilenvektoren in B mit $ACB^T = 0$ und $BCA^T = (ACB^T)^T$ verwendet haben. Die Kovarianzmatrix von Z ist daher blockdiagonal. Die Determinante einer Blockmatrix ist das Produkt der Determinante der Blöcke d.h.

$$\det \begin{pmatrix} ACA^T & 0 \\ 0 & BCB^T \end{pmatrix} = \det(ACA^T)\, \det(BCB^T).$$

Damit hat Z die Dichte mit $z = (y, u)$ mit $x \in \mathbb{R}^m$ und $u \in \mathbb{R}^{n-m}$,

$$\begin{aligned}
f_Z(z) &= e^{-\frac{1}{2}\left\langle (z-c-T\mu), (TCT^T)^{-1}(z-c-T\mu)\right\rangle} \frac{1}{\sqrt{\det(TCT^T)}} \frac{1}{(2\pi)^{\frac{n}{2}}} \\
&= e^{-\frac{1}{2}\left\langle (y-b-A\mu), (ACA^T)^{-1}(y-b-A\mu)\right\rangle} e^{-\frac{1}{2}\left\langle (u-B\mu), (BCB^T)^{-1}(u-B\mu)\right\rangle} \frac{1}{\sqrt{\det(ACA^T)}\sqrt{\det(BCB^T)}} \frac{1}{(2\pi)^{\frac{n}{2}}} \\
&= \frac{1}{\sqrt{(2\pi)^m \det(ACA^T)}} e^{-\frac{1}{2}\left\langle (y-b-A\mu), (ACA^T)^{-1}(y-b-A\mu)\right\rangle} \frac{1}{\sqrt{(2\pi)^{n-m} \det(BCB^T)}} e^{-\frac{1}{2}\left\langle (u-B\mu), (BCB^T)^{-1}(u-B\mu)\right\rangle} \\
&= f_Y(y)\, f_U(u)
\end{aligned}$$

Damit sind $Y = AX + b$ (die ersten m Komponenten von Z) und $U = BX$ (die letzten $n - m$ Komponenten von Z) unabhängige Zufallsvariablen, da die gemeinsame Dichte f_Z in die Dichte f_Y von Y und f_U von U faktorisiert. Damit ergibt sich sofort die Verteilung von Y,

$$Y \sim \mathcal{N}(A\mu + b,\, ACA^T).$$

Im allgemeinen würde man die gemeinsame Dichte der ersten m Komponenten als Randverteilung der gemeinsamen Dichte von Z erhalten, d.h.

$$\begin{aligned}
f_Y(y_1, \dots, y_m) &= \int_{\mathbb{R}^{n-m}} f_Z(y_1, \dots, y_m, z_{m+1}, \dots, z_n)\, dz_{m+1} \dots dz_n \\
&= \int_{\mathbb{R}^{n-m}} f_Y(y_1, \dots, y_m) f_U(z_{m+1}, \dots, z_n)\, dz_{m+1} \dots dz_n \\
&= f_Y(y_1, \dots, y_m).
\end{aligned}$$

Remark 0.1 *Anwendung: Wir bestimmen die Verteilung einer Summe* $Z = \sum_{i=1}^n X_i$ *von unabhängigen Zufallsvariablen, wobei* $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ *(d.h.* ***nicht*** *identisch verteilt). Sei* $\mathrm{e} \in \mathbb{R}^n$ *der Vektor bei dem jede Komponente eins ist. Es gilt* $Z = \mathrm{e}^T X$ *wobei* $X = (X_1, \ldots, X_n)$ *und* $X \sim \mathcal{N}(\mu, C)$ *mit* $C_{ii} = \sigma_i^2$ *und* $C_{ij} = 0$ *für* $i \neq j$*, da* X_i *unabhängig ist von* X_j*. Anwendung des Theorems 0.2 liefert* $Z \sim \mathcal{N}(\mathrm{e}^T \mu, \mathrm{e}^T C \, \mathrm{e}) = \mathcal{N}(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2)$*. Damit erhalten wir das Resultat der Übungsaufgabe deutlich schneller (wer das so löst bekommt natürlich keine Punkte) - allerdings mussten wir dazu erst den etwas technischen Beweis führen.*

□

Das Resultat des nächsten Lemmas wurde schon implizit im Beweis des vorigen Satzes benützt.

Lemma 0.1 *Sei* $X \sim \mathcal{N}(\mu, C)$ *mit* $\mu \in \mathbb{R}^n$ *und* $C \in \mathbb{R}^{n \times n}$*. Sei* $A \in \mathbb{R}^{r \times n}$ *und* $B \in \mathbb{R}^{s \times n}$*. Dann sind* AX *und* BX *unabhängig genau dann wenn* $ACB^T = 0$*.*

Proof: Mittels des Theorems 0.2 erhalten wir die Verteilung von AX als $\mathcal{N}(A\mu, ACA^T)$ und von BX als $\mathcal{N}(B\mu, BCB^T)$. Für die gemeinsame Verteilung von (AX, BX) erhält man, $\mathcal{N}((A\mu, B\mu), \Sigma)$ wobei die Kovarianzmatrix Σ gegeben ist als

$$\Sigma = \begin{pmatrix} ACA^T & ACB^T \\ BCA^T & BCB^T \end{pmatrix}.$$

Sind AX und BX unabhängig, dann ist deren Kovarianz null d.h. dann gilt $ACB^T = 0$ und $BCA^T = (ACB^T)^T = 0$. Umgekehrt folgt aus $ACB^T = 0$ die blockdiagonale Struktur der Kovarianzmatrix von (AX, BX). Damit faktorisiert aber die Dichtefunktion der gemeinsamen Gaussverteilung in das Produkt der Dichten von AX und BX d.h. sie sind unabhängig. □

1 Die χ^2-Verteilung

Definition 1.1 *Eine Zufallsvariable* X *ist* χ^2*-verteilt mit Parameter* m *oder kurz* χ_m^2 *verteilt, wenn sie folgende Dichte besitzt*

$$f(x) = \begin{cases} 0, & \text{if } x \leq 0, \\ K_m x^{\frac{m-2}{2}} \, e^{-\frac{x}{2}}, & \text{if } x > 0. \end{cases}, \quad \text{where} \quad K_m = \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)},$$

wobei $\Gamma(x)$ *die Gamma-Funktion ist.*

Die wichtigste Eigenschaft der χ^2-Verteilung ist der Zusammenhang mit der Gauss-Verteilung.

Proposition 1.1 *Seien* $Z_1, \ldots, Z_m$ *unabhängige und identisch verteilte Zufallsvariablen mit* $Z_i \sim \mathcal{N}(0,1)$ *für* $i = 1, \ldots, m$*. Dann hat*

$$X = \sum_{i=1}^m Z_i^2,$$

eine χ_m^2*-Verteilung. Darüberhinaus gilt für* $X \sim \chi_k^2$ *and* $Y \sim \chi_l^2$*, wobei* X *und* Y *unabhängig sind,*

$$X + Y \sim \chi_{k+l}^2.$$

Proof: Es gilt $X = \sum_{i=1}^m Z_i^2$ mit $Z \in \mathcal{N}(0_m, \mathbb{1}_m)$ und daher

$$\mathrm{P}(X \leq x) = \mathrm{P}\Big(\sum_{i=1}^m Z_i^2 \leq x\Big) = \frac{1}{(2\pi)^{\frac{m}{2}}} \int_{\sum_{i=1}^n z_i^2 \leq x} e^{-\frac{1}{2}\sum_{i=1}^m z_i^2} \, dz_1 \ldots dz_m.$$

Wir führen m-dimensionale Kugelkoordinaten ein und erhalten dadurch den Faktor r^{m-1}. Da die zu integrierende Funktion nur vom Radius r abhängt ($r^2 = \sum_{i=1}^m z_i^2$) können wir den Teil, der die

Oberfläche parametrisiert, direkt herausintegrieren:

$$
\begin{aligned}
\frac{1}{(2\pi)^{\frac{m}{2}}} \int_{\sum_{i=1}^n z_i^2 \le x} e^{-\frac{1}{2}\sum_{i=1}^m z_i^2}\, dz_1 \dots dz_m &= \frac{1}{(2\pi)^{\frac{m}{2}}} \int_{r^2 \le x} e^{-\frac{r^2}{2}} r^{m-1}\, dr\, d\Omega \\
&= \frac{2\pi^{\frac{m}{2}}}{(2\pi)^{\frac{m}{2}}\, \Gamma\left(\frac{m}{2}\right)} \int_0^{\sqrt{x}} e^{-\frac{r^2}{2}} r^{m-1}\, dr \\
&= \frac{1}{2^{\frac{m}{2}} \Gamma\left(\frac{m}{2}\right)} \int_0^x e^{-\frac{u}{2}}\, u^{\frac{m}{2}-1}\, du,
\end{aligned}
$$

wobei Ω die Parametrisierung der Oberfläche der m-dimensionalen Einheitskugel darstellt. Die Integration über die Oberfläche ergibt dann die Kugeloberfläche O_m der Einheitskugel in m Dimensionen

$$
O_m = \frac{2\pi^{\frac{m}{2}}}{\Gamma\left(\frac{m}{2}\right)}.
$$

Der letzte Schritt ergibt sich dann mit der Variablentransformation $u = r^2$. Der Hauptsatz der Differential- und Integralrechnung liefert dann die Dichtefunktion als Ableitung der Verteilungsfunktion $F_X(x) = \mathrm{P}(X \le x)$.

Das zweite Resultat ergibt sich direkt aus dem ersten. Es folgt $X = \sum_{i=1}^k Z_i^2$ mit $Z_i \sim \mathcal{N}(0,1)$ und $Y = \sum_{j=1}^l U_j^2$ mit $U_j \in \mathcal{N}(0,1)$. Mit der Unabhängigkeit von X und Y folgt dann,

$$
X + Y = \sum_{i=1}^{k+l} Z_i^2 \quad \sim \quad \chi^2_{k+l},
$$

wobei $Z_i \sim \mathcal{N}(0,1)$ für $i = 1, \dots, k+l$. □

2 Die Student t-Verteilung

Definition 2.1 *Eine Zufallsvariable X ist Student t-verteilt mit Paramter m oder kurz t_m-verteilt, wenn sie folgende Dichte besitzt:*

$$
f(x) = L_m \left(1 + \frac{x^2}{m}\right)^{-\frac{m+1}{2}}, \quad \textit{wobei} \quad L_m = \frac{\Gamma\left(\frac{m+1}{2}\right)}{\sqrt{\pi\, m}\, \Gamma\left(\frac{m}{2}\right)}.
$$

Wie bei der χ^2-Verteilung ergeben sich die Hauptanwendungen aus folgendem Resultat.

Proposition 2.1 *Seien $Z \sim \mathcal{N}(0,1)$ und $U \sim \chi^2_m$ zwei unabhängige Zufallsvariablen. Dann ist $\frac{Z}{\sqrt{\frac{U}{m}}}$ t_m-verteilt.*

Proof: Wir betrachten die Transformation, $\phi : \mathbb{R} \times \mathbb{R}_+ \to \mathbb{R} \times \mathbb{R}_+$,

$$
\begin{pmatrix} x \\ y \end{pmatrix} = \phi(z,u) = \begin{pmatrix} \frac{z}{\sqrt{\frac{u}{m}}} \\ u \end{pmatrix}.
$$

Die Umkehrabbildung ϕ^{-1} existiert und ist

$$
\begin{pmatrix} z \\ u \end{pmatrix} = \phi^{-1}(x,y) = \begin{pmatrix} x\sqrt{\frac{y}{m}} \\ y \end{pmatrix}.
$$

Die Jacobi-matrix $J_{\phi^{-1}}(x,y)$ ist

$$J_{\phi^{-1}}(x,y) = \begin{pmatrix} \sqrt{\frac{y}{m}} & \frac{x}{2}\sqrt{\frac{1}{m\,y}} \\ 0 & 1 \end{pmatrix}$$

mit Determinante $\det J_{\phi^{-1}}(x,y) = \sqrt{\frac{y}{m}}$. Damit ergibt sich die gemeinsame Dichte von $(X,Y) = \phi(Z,U)$ unter Verwendung der Unabhängigkeit von Z und U als

$$\begin{aligned} f(x,y) &= f_Z(\phi^{-1}(x,y))\, f_U(\phi^{-1}(x,y))\, |\det J_{\phi^{-1}(x,y)}| \\ &= \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2 y}{2m}} \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} y^{\frac{m-2}{2}}\, e^{-\frac{y}{2}} \sqrt{\frac{y}{m}} \end{aligned}$$

Die gesuchte Dichte der ersten Komponente berechnen wir als Randverteilung der gemeinsamen Dichte:

$$\begin{aligned} f(x) &= \int_0^\infty f(x,y)\,dy \\ &= \int_0^\infty \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2 y}{2m}} \frac{1}{2^{\frac{m}{2}}\Gamma(\frac{m}{2})} y^{\frac{m-2}{2}}\, e^{-\frac{y}{2}} \sqrt{\frac{y}{m}}, dy \\ &= \frac{1}{\sqrt{\pi m}\,\Gamma(\frac{m}{2})} \frac{1}{2^{\frac{m+1}{2}}} \int_0^\infty y^{\frac{m-1}{2}} e^{-y\left(\frac{1+\frac{x^2}{m}}{2}\right)}\, dy \end{aligned}$$

Wir führen die Variablentransformation $w = \left(\frac{1+\frac{x^2}{m}}{2}\right)y$ durch und erhalten

$$\begin{aligned} f(x) &= \frac{1}{\sqrt{\pi m}\,\Gamma(\frac{m}{2})} \frac{1}{2^{\frac{m+1}{2}}} \int_0^\infty \left(\frac{2w}{(1+\frac{x^2}{m})}\right)^{\frac{m-1}{2}} e^{-w} \frac{2}{(1+\frac{x^2}{m})}\, dw \\ &= \frac{1}{\sqrt{\pi m}\,\Gamma(\frac{m}{2})} \left(1+\frac{x^2}{m}\right)^{-\frac{m+1}{2}} \int_0^\infty w^{\frac{m+1}{2}-1} e^{-w}\, dw \\ &= \frac{\Gamma(\frac{m+1}{2})}{\sqrt{\pi m}\,\Gamma(\frac{m}{2})} \left(1+\frac{x^2}{m}\right)^{-\frac{m+1}{2}} \end{aligned}$$

Im letzten Schritt haben wir die Definition der Gammafunktion, $\Gamma(x) = \int_0^\infty t^{x-1} e^{-t}\, dt$, verwendet. Damit erhalten wir das gewünschte Ergebnis. □

3 Verteilung der Schätzer von Mittelwert und Varianz bei normalverteilter Stichprobe

Bevor wir zur Verteilung der Schätzer kommen noch folgender wichtiger Satz.

Proposition 3.1 *Seien X und Y unabhängige Zufallsvariablen und $f, g : \mathbb{R} \to \mathbb{R}$. Dann sind auch $f(X)$ und $g(Y)$ unabhängige Zufallsvariablen.*

Proof: Wir beweisen diesen Satz nur für den Fall, daß f, g und deren Umkehrabbildungen stetig diferenzierbar sind. Sei $(u,v) = \phi(x,y) = (f(x), g(y))$, dann ist $(x,y) = \phi^{-1}(u,v) = (f^{-1}(u), g^{-1}(v))$. Man beachte, daß die Jacobimatrix nur Einträge auf der Diagonalen hat und daher $\det J_{\phi^{-1}}(u,v) = \frac{\partial f^{-1}(u)}{\partial u}\frac{\partial g^{-1}(v)}{\partial v}$. Unter Ausnützung der Unabhängigkeit von X und Y (d.h. die gemeinsame Dichte faktorisiert: $f_{X,Y}(x,y) = f_X(x)\, f_Y(y)$) gilt:

$$\begin{aligned} f(u,v) &= f_X(\phi^{-1}(u,v))\, f_Y(\phi^{-1}(u,v))\, |\det \phi^{-1}(u,v)| \\ &= f_X(f^{-1}(u))\, f_Y(g^{-1}(v))\, \frac{\partial f^{-1}(u)}{\partial u} \frac{\partial g^{-1}(v)}{\partial v} \\ &= \left(f_X(f^{-1}(u))\, \frac{\partial f^{-1}(u)}{\partial u}\right)\left(f_Y(g^{-1}(v))\, \frac{\partial g^{-1}(v)}{\partial v}\right) \end{aligned}$$

Die gemeinsame Dichte von $U = f(X)$ und $V = g(Y)$ faktorisiert in ein Produkt. Daher sind $f(X)$ und $g(Y)$ unabhängig. $\square$

Proposition 3.2 *Seien $(X_1, \dots, X_n)$ unabhängige und identisch verteilte Zufallsvariablen mit $X_i \sim \mathcal{N}(\mu, \sigma^2)$. Wir bezeichnen mit $\hat{\mu} = \frac{1}{n}\sum_{i=1}^n X_i$ den Schätzer des Mittelwerts μ und mit $\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \hat{\mu})^2$ den Schätzer der Varianz σ^2. Es gilt*

- *$\hat{\mu}$ und $\hat{\sigma}^2$ sind unabhängig,*
- *$\frac{n-1}{\sigma^2}\hat{\sigma}^2$ hat eine χ^2_{n-1}-Verteilung,*
- *$\frac{\mu-\hat{\mu}}{\sqrt{\frac{\hat{\sigma}^2}{n}}}$ hat eine t_{n-1}-Verteilung.*

Proof: Wir definieren die **Zentrierungsmatrix** $Z = \mathbb{1}_n - \frac{1}{n}\,\mathrm{e}\,\mathrm{e}^T$ und fassen die Stichprobe zusammen als $X = (X_1, \dots, X_n)^T$. Dann kann der Varianz-Schätzer $\hat{\sigma}^2$ geschrieben werden als

$$\hat{\sigma}^2 = \frac{1}{n-1}(ZX)^T(ZX) = \frac{1}{n-1}X^T ZX,$$

wobei wir ausgenützt haben, daß $ZZ^T = Z$ (Z ist eine Projektionsmatrix d.h. $Z = Z^T$ und $ZZ = Z$). Den Mittelwertschätzer schreiben wir als $\hat{\mu} = \frac{1}{n}\,\mathrm{e}^T X$. Da $X \sim \mathcal{N}(\mu, \sigma^2 \mathbb{1}_n)$ und sowohl ZX als auch $\mathrm{e}^T X$ lineare Funktionen von einer Gaussvariable sind, können wir Lemma 0.1 anwenden. Danach sind ZX und $\mathrm{e}^T X$ unabhängig, wenn $Z\mathbb{1}_m\,\mathrm{e} = 0$. Wir rechnen nach

$$Z\mathbb{1}_m\,\mathrm{e} = Z\,\mathrm{e} = \mathrm{e} - \frac{1}{n}\,\mathrm{e}(\mathrm{e}^T\,\mathrm{e}) = \mathrm{e} - \mathrm{e} = 0.$$

Da nach Proposition 3.1 Funktionen von unabhängigen Zufallsvariablen wieder unabhängig sind, sind auch $\hat{\sigma}^2 = \frac{1}{n-1}\|ZX\|^2$ und $\hat{\mu} = \frac{1}{n}\,\mathrm{e}^T X$ unabhängig.

Für die Herleitung des zweiten Resultats schreiben wir $\hat{\sigma}^2$ mittels standardnormalverteilter Zufallsvariablen um:

$$X^* = \frac{1}{\sigma}(X - \mu\ \mathrm{e}) \quad\Longrightarrow\quad X^* \sim \mathcal{N}(0_n, \mathbb{1}_n),$$

wobei 0_n der n-dimensionale Nullvektor ist. Wir erhalten

$$\hat{\sigma}^2 = \frac{1}{n-1}(\sigma X^* + \mu\ \mathrm{e})^T Z(\sigma X^* + \mu\ \mathrm{e}) = \frac{\sigma^2}{n-1}(X^*)^T ZX^*.$$

Wir definieren eine orthogonale Matrix A, wobei wir die erste Zeile als $\frac{1}{\sqrt{n}}\,\mathrm{e}$ definieren. So eine Matrix kann immer mit Hilfe des Gram-Schmidt Orthogonalisierungsverfahren erzeugt werden. Wir definieren

$$Y^* = AX^* \quad\Longrightarrow\quad X^* = A^T Y^*,$$

wobei wir die Eigenschaft $A^{-1} = A^T$ einer orthogonalen Matrix ausgenützt haben. Y^* hat eine $\mathcal{N}(A0_n, A\mathbb{1}_n A^T) = \mathcal{N}(0_n, \mathbb{1}_n)$-Verteilung. Es gilt:

$$\begin{aligned}
\hat{\sigma}^2 &= \frac{\sigma^2}{n-1}\,(A^T Y^*)^T Z A^T Y^* = \sigma^2 (Y^*)^T AZA^T Y^* \\
&= \frac{\sigma^2}{n-1}\,(Y^*)^T A\Big(\mathbb{1}_n - \frac{1}{n}\,\mathrm{e}\,\mathrm{e}^T\Big)A^T Y^* \\
&= \frac{\sigma^2}{n-1}\,Y^*\Big(\mathbb{1}_n - \frac{1}{n}A\,\mathrm{e}\,\mathrm{e}^T A^T\Big)Y^* = \frac{\sigma^2}{n-1}\sum_{i=\mathbf{2}}^n (Y_i^*)^2,
\end{aligned}$$

wobei wir verwendet haben, daß $A\,\mathrm{e} = \begin{pmatrix}\sqrt{n}\\ 0\\ \vdots\\ 0\end{pmatrix}$. Mittels Proposition 1.1 folgt, daß $\frac{n-1}{\sigma^2}\hat{\sigma}^2 \sim \chi^2_{n-1}$.

Wir führen ein:

$$Z = \sqrt{\frac{n}{\sigma}}(\mu - \hat{\mu}), \quad X = \frac{(n-1)}{\sigma^2}\hat{\sigma}^2.$$

Wie in der Vorlesung gezeigt, gilt $Z \sim \mathcal{N}(0,1)$ und X hat mit Hilfe des zweiten Resultats eine χ^2_{n-1}-Verteilung. Wir erhalten:

$$\frac{\mu - \hat{\mu}}{\sqrt{\frac{\hat{\sigma}^2}{n}}} = \frac{\sqrt{n}(\mu - \hat{\mu})}{\sqrt{\sigma^2}} \frac{1}{\sqrt{\frac{(n-1)\hat{\sigma}^2}{(n-1)\sigma^2}}} = \frac{Z}{\sqrt{\frac{X}{n-1}}}.$$

Die Anwendung von Proposition 2.1 liefert das Ergebnis. $\square$