Abhängigkeitsmaße

Seien X_1 und X_2 zwei Zufallsvariablen. Es gibt einige skalare Maße für die Abhängigkeit zwischen X_1 und X_2.

Lineare Korrelation

Annahme: $var(X_1), var(X_2) \in (0, \infty)$. Der Koeffizient der linearen Korrelation $\rho_L(X_1, X_2)$ ist folgendermaßen gegeben:

$$\rho_L(X_1, X_2) = \frac{cov(X_1, X_2)}{\sqrt{var(X_1)var(X_2)}}$$

$$X_1 \text{ und } X_2 \text{ sind unabhängig } \Rightarrow \rho_L(X_1, X_2) = 0$$

$$\rho_L(X_1, X_2) = 0 \text{ impliziert nicht, dass } X_1 \text{ und } X_2 \text{ unabhängig sind}$$

Beispiel 2 *Sei $X_1 \sim N(0,1)$ und $X_2 = X_1^2$. Es gilt $\rho_L(X_1, X_2) = 0$ aber X_1 und X_2 sind klarerweise abhängig.*

Weiters gilt:

$$|\rho_L(X_1, X_2)| = 1 \Leftrightarrow \exists \alpha, \beta \in \mathbb{R},\ \beta \neq 0,\ \text{sodass } X_2 \stackrel{d}{=} \alpha + \beta X_1$$

$$\text{und } \text{signum}(\beta) = \text{signum}(\rho_L(X_1, X_2))$$

Der lineare Korrelationskoeff. ist eine Invariante unter streng monoton steigende lineare Transformationen. D.h. für zwei Zufallsvariablen X_1 und X_2 und reellen Konstanten $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$, $\beta_1 > 0$ und $\beta_2 > 0$ gilt:

$$\rho_L(\alpha_1 + \beta_1 X_1, \alpha_2 + \beta_2 X_2) = \rho_L(X_1, X_2).$$

Der lineare Korrelationskoeffizient ist jedoch keine Invariante unter streng monoton steigende nicht-lineare Transformationen.

Beispiel 3 *Seien $X \sim Exp(\lambda)$, $X_2 = X_1$, und T_1, T_2 zwei streng monoton steigende Transformationen: $T_1(X_1) = X_1$ und $T_2(X_1)) = X_1^2$. Dann gilt:*

$$\rho_L(X_1, X_1) = 1 \textit{ und } \rho_L(T_1(X_1), T_2(X_1)) = \frac{2}{\sqrt{5}}.$$

Rang Korrelation

Die Koeffizienten der Rang Korrelation (Spearmans Rho und Kendalls Tau) sind Maße für die Übereinstimmung von bivariaten Zufallsvektoren.

Seien (x_1, x_2) und $(\tilde{x}_1, \tilde{x}_2)$ zwei Punkte in $\mathbb{R}^2$. Die zwei Punkte heißen *übereinstimmend* wenn $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) > 0$ und *nicht übereinstimmend* wenn $(x_1 - \tilde{x}_1)(x_2 - \tilde{x}_2) < 0$.

Seien $(X_1, X_2)^T$ und $(\tilde{X}_1, \tilde{X}_2)^T$ zwei unabhängige Zufallsvektoren mit identischer bivariater Verteilung.

Die Kendall's Tau ρ_τ ist definiert als

$$\rho_\tau(X_1, X_2) = P\left((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) > 0\right) - P\left((X_1 - \tilde{X}_1)(X_2 - \tilde{X}_2) < 0\right)$$

Sei $(\hat{X}_1, \hat{X}_2)$ ein dritter von (X_1, X_2) und $(\tilde{X}_1, \tilde{X}_2)$ unahängiger Zufallsvektor mit derselben Verteilung wie (X_1, X_2) und $(\tilde{X}_1, \tilde{X}_2)$.

Die Spearman's Rho ρ_S ist definiert als

$$\rho_S(X_1, X_2) = 3\left\{P\left((X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) > 0\right) - P\left((X_1 - \tilde{X}_1)(X_2 - \hat{X}_2) < 0\right)\right\}$$

Einige Eigenschaften von ρ_τ und ρ_S:

- $\rho_\tau(X_1, X_2) \in [-1, 1]$ und $\rho_S(X_1, X_2) \in [-1, 1]$.

- Wenn X_1 und X_2 unabhängig, dann $\rho_\tau(X_1, X_2) = \rho_S(X_1, X_2) = 0$. Die Umkehrung gilt i.a. nicht.

- Sei $T\colon \mathbb{R} \to \mathbb{R}$ eine streng monoton steigende Funktion. Dann gilt:

$$\rho_\tau(T(X_1), T(X_2)) = \rho_\tau(X_1, X_2)$$

$$\rho_S(T(X_1), T(X_2)) = \rho_S(X_1, X_2)$$

Beweis: 1) und 2) sind trivial. Beweis von 3) erfolgt mit Hilfe von Copulas, später...

Tail-Abhängigkeit

Definition 1 *Sei $(X_1, X_2)^T$ ein Zufallsvektor mit Randverteilungen F_1 und F_2. Der Koeffizient der oberen Tail-Abhängigkeit von $(X_1, X_2)^T$ wird folgendermaßen definiert:*

$$\lambda_U(X_1, X_2) = \lim_{u \to 1^-} P(X_2 > F_2^{\leftarrow}(u) | X_1 > F_1^{\leftarrow}(u))$$

vorausgesetzt der Limes existiert.

Der Koeffizient der unteren Tail-Abhängigkeit von $(X_1, X_2)^T$ wird folgendermaßen definiert:

$$\lambda_L(X_1, X_2) = \lim_{u \to 0^+} P(X_2 \leq F_2^{\leftarrow}(u) | X_1 \leq F_1^{\leftarrow}(u))$$

vorausgesetzt der Limes existiert.

Wenn $\lambda_U > 0$ ($\lambda_L > 0$) heißt es, $(X_1, X_2)^T$ hat eine obere (untere) Tail-Abhängigkeit.

(Siehe Joe 1997, Schmidt und Stadtmüller 2002)

Übung 1 *Sei $X_1 \sim Exp(\lambda)$ und $X_2 = X_1^2$. Bestimmen Sie $\lambda_U(X_1, X_2)$, $\lambda_L(X_1, X_2)$ und zeigen Sie, dass $(X_1, X_2)^T$ eine obere und eine untere Tail-Abhängigkeit hat. Berechnen Sie auch den linearen Korrelationskoeffizienten $\rho_L(X_1, X_2)$.*

Multivariate elliptische Verteilungen

a) Die multivariate Normalverteilung

Definition 2 *Der Zufallsvektor* $(X_1, X_2, \ldots, X_d)^T$ *hat eine multivariate Normalverteilung (oder eine multivariate Gauss'sche Verteilung) wenn* $X \stackrel{d}{=} \mu + AZ$, *wobei* $Z = (Z_1, Z_2, \ldots, Z_k)^T$ *ein Vektor von i.i.d. normalverteilten ZV* ($Z_i \sim N(0,1)$, $\forall i = 1, 2, \ldots, k$), $A \in \mathbb{R}^{d \times k}$ *ist eine konstante Matrix und* $\mu \in \mathbb{R}^d$ *ist ein konstanter Vektor.*

Für so einen Zufallsvektor X gilt: $E(X) = \mu$, $cov(X) = \Sigma = AA^T$ (Σ positiv semidefinit). Notation: $X \sim N_d(\mu, \Sigma)$.

Theorem 1 *(Multivariate Normalverteilung: äquiv. Definitionen)*

1. $X \sim N_d(\mu, \Sigma)$ *für einen Vektor* $\mu \in \mathbb{R}^d$ *und eine positiv semidefinite Matrix* $\Sigma \in \mathbb{R}^{d \times d}$, *dann und nur dann wenn* $\forall a \in \mathbb{R}^d$, $a = (a_1, a_2, \ldots, a_d)^T$, *die Zufallsvariable* $a^T X$ *normal verteilt ist.*

2. *Ein Zufallsvektor* $X \in \mathbb{R}^d$ *ist multivariat normal verteilt dann und nur dann wenn seine charakteristische Funktion folgendermaßen gegeben ist:*

$$\phi_X(t) = E(\exp\{it^T X\}) = \exp\{it^T \mu - \frac{1}{2} t^T \Sigma t\}$$

für einen Vektor $\mu \in \mathbb{R}^d$ *und eine positiv semidefinite Matrix* $\Sigma \in \mathbb{R}^{d \times d}$.

3. Ein Zufallsvektor $X \in \mathbb{R}^d$ mit $E(X) = \mu$ und $cov(X) = \Sigma$, wobei die Determinante von Σ positiv ist ($det(\Sigma) := |\Sigma| > 0$), ist normal verteilt, d.h. $X \sim N_d(\mu, \Sigma)$, dann und nur dann wenn seine Dichtefunktion folgendermaßen gegeben ist

$$f_X(x) = \frac{1}{\sqrt{(2\pi)^d|\Sigma|}} \exp\left\{-\frac{(x-\mu)^T\Sigma^{-1}(x-\mu)}{2}\right\}.$$

Beweis: (siehe zB. Gut 1995)

Theorem 2 *(Eigenschaften der multivariaten Normalverteilung)* *Für $X \sim N_d(\mu, \Sigma)$ gilt:*

Lineare Kombinationen:
Für $B \in \mathbb{R}^{k\times d}$ und $b \in \mathbb{R}^k$. Es gilt dann $BX+b \in N_k(B\mu+b, B\Sigma B^T)$.

Randverteilungen:
Setze $X^T = \left(X^{(1)^T}, X^{(2)^T}\right)$ für $X^{(1)^T} = (X_1, X_2, \ldots, X_k)^T$ und $X^{(2)^T} = (X_{k+1}, X_{k+2}, \ldots, X_d)^T$ und analog

$$\mu^T = \left(\mu^{(1)^T}, \mu^{(2)^T}\right) \text{ und } \Sigma = \begin{pmatrix} \Sigma^{(1,1)} & \Sigma^{(1,2)} \\ \Sigma^{(2,1)} & \Sigma^{(2,2)} \end{pmatrix}.$$

Es gilt dann $X^{(1)} \sim N_k\left(\mu^{(1)}, \Sigma^{(1,1)}\right)$ und $X^{(2)} \sim N_{d-k}\left(\mu^{(2)}, \Sigma^{(2,2)}\right)$.

Bedingte Verteilungen:

Wenn Σ regulär, dann ist auch der bedingte Z.Vektor $X^{(2)} \Big| X^{(1)} = x^{(1)}$ multivariat normal verteilt

$$X^{(2)} | X^{(1)} = x^{(1)} \sim N_{d-k}\Big(\mu^{(2,1)}, \Sigma^{(22,1)}\Big) \text{ wobei}$$

$$\mu^{(2,1)} = \mu^{(2)} + \Sigma^{(2,1)}\Big(\Sigma^{(1,1)}\Big)^{-1}\Big(x^{(1)} - \mu^{(1)}\Big) \text{ und}$$

$$\Sigma^{(22,1)} = \Sigma^{(2,2)} - \Sigma^{(2,1)}\Big(\Sigma^{(1,1)}\Big)^{-1}\Sigma^{(1,2)}.$$

Quadratische Formen:

Wenn Σ regulär, dann gilt $D^2 = (X - \mu)^T \Sigma^{-1} (X - \mu) \sim \chi^2_d$.
Die Zufallsvariable D heißt *Mahalanobis Distanz*.

Faltung:

Seien $X \sim N_d(\mu, \Sigma)$ und $Y \sim N_d(\tilde{\mu}, \tilde{\Sigma})$ zwei unabhängige Zufallsvektoren. Es gilt dann $X + Y \sim N_d(\mu + \tilde{\mu}, \Sigma + \tilde{\Sigma})$.

b) Varianz-gemischte Normalverteilungen

Definition 3 *Ein Zufallsvektor $X \in \mathbb{R}^d$ hat eine multivariate Varianz-gemischte Normalverteilung wenn $X \stackrel{d}{=} \mu + WAZ$ wobei: $Z \sim N_k(0, I)$, $W \geq 0$ ist eine von Z unabhängige positive Zufallsvariable, $\mu \in \mathbb{R}^d$ ist ein konstanter Vektor, $A \in \mathbb{R}^{d \times k}$ ist eine konstante Matrix, und I ist die Einheitsmatrix.*

Unter der Bedingung $W = w$ ist X normalverteilt: $X \sim N_d(\mu, w^2\Sigma)$, wobei $\Sigma = AA^T$.

$E(X) = \mu$ und $cov(X) = E(W^2AZZ^TA^T) = E(W^2)\Sigma$ falls $E(W^2) < \infty$

Beispiel 4 Die multivariate t_α-Verteilung
Sei $Y \sim IG(\alpha, \beta)$ (Inverse Gamma-Verteilung) mit Dichtefunktion:

$$f_{\alpha,\beta}(x) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{-(\alpha+1)} \exp(-\beta/x) \;\; x > 0, \alpha > 0, \beta > 0$$

Dann gilt:

$$E(Y) = \frac{\beta}{\alpha - 1} \text{ für } \alpha > 1, \;\; var(Y) = \frac{\beta^2}{(\alpha-1)^2(\alpha-2)} \text{ für } \alpha > 2$$

Sei $W^2 \sim IG(\alpha/2, \alpha/2)$. Dann ist die Verteilung von $X = \mu + WAZ$ eine multivariate t_α Verteilung mit α Freiheitsgraden: $X \sim t_d(\alpha, \mu, \Sigma)$.

$$cov(X) = E(W^2)\Sigma = \frac{\alpha}{\alpha - 2}\Sigma$$

c) Sphärische Verteilungen

Definition 4 *Ein Zufallsvektor* $X = (X_1, X_2, \ldots, X_d)^T$ *hat eine sphärische Verteilung wenn für jede orthogonale Matrix* $U \in \mathbb{R}^{d\times d}$ *die Gleichung* $UX \stackrel{d}{=} X$ *gilt.*

Theorem 3 *Die folgenden Aussagen sind äquivalent.*

1. *Der Zufallsvektor* $X \in \mathbb{R}^d$ *hat eine sphärische Verteilung.*

2. *Es existiert eine Funktion* $\psi \colon \mathbb{R} \to \mathbb{R}$*, sodass die charakteristische Funktion von* X *folgendermaßen gegeben wird:*
$$\phi_X(t) = \psi(t^T t) = \psi(t_1^2 + t_2^2 + \ldots + t_d^2)$$

3. *Für jeden Vektor* $a \in \mathbb{R}^d$ *gilt* $a^t X \stackrel{d}{=} ||a|| X_1$ *wobei* $||a||^2 = a_1^2 + a_2^2 + \ldots + a_d^2$.

4. X *lässt sich als* $X \stackrel{d}{=} RS$ *repräsentieren, wobei der Zufallsvektor* $S \in \mathbb{R}^d$ *gleichmäßig verteilt auf der Einheitskugel* S^{d-1}*,* $\mathcal{S}^{d-1} = \{x \in \mathbb{R}^d \colon ||x|| = 1\}$*, ist, und* $R \geq 0$ *eine von* S *unabhängige ZV ist.*

Notation einer sphärischen Verteilung: $X \sim S_d(\psi)$

Beispiel 5 *Die standard Normalverteilung ist eine sphärische Verteilungen.*

Sei $X \sim N_d(0, I)$. *Dann* $X \sim S_d(\psi)$ *mit* $\psi = \exp(-x/2)$.

Tatsächlich: $\phi_X(t) = \exp\{it^T 0 - \frac{1}{2} t^T I t\} = \exp\{-t^T t/2\} = \psi(t^T t)$.

Sei $X = RS$ *die stochastische Darstellung von* $X \sim N_d(0, I)$. *Es gilt* $||X||^2 \stackrel{d}{=} R^2 \sim \chi^2_d$;

Simulation einer sphärischen Verteilung:

(i) Simuliere s aus einer gleichmäßig verteilten Zufallsvektor in $\mathcal{S}^{d-1}$ (zB. in dem y aus einer multivariaten Standard Normalverteilung $Y \sim N_d(0, I)$ simuliert und $s = y/||y||$ gesetzt wird).

(ii) Simuliere r aus R.

(iii) Setze $x = rs$.

d) Elliptische Verteilungen

Definition 5 *Ein Zufallsvektor $X \in \mathbb{R}^d$ hat eine elliptische Verteilung wenn $X \stackrel{d}{=} \mu + AY$, wobei $Y \sim S_k(\psi)$, $\mu \in \mathbb{R}^d$ ist ein konstanter Vektor und $A \in \mathbb{R}^{d\times k}$ ist eine konstante Matrix.*

Die charakteristische Funktion:

$$\phi_X(t) = E(\exp\{it^T X\}) = E(\exp\{it^T(\mu + AY)\}) = \exp\{it^T\mu\}E(\exp\{i(A^T t)^T Y\})$$

$$= \exp\{it^T\mu\}\psi(t^T\Sigma t),$$

wobei $\Sigma = AA^T$.

Notation elliptische Verteilungen: $X \sim E_d(\mu, \Sigma, \psi)$
μ heißt Positionsparameter (*location parameter*),
Σ heißt Dispersionsparameter (*dispersion parameter*),
ψ heißt charakteristischer Generator der elliptischen Verteilung.

Falls $A \in \mathbb{R}^{d\times d}$ regulär, dann gilt folgende Relation zwischen elliptischen und sphärischen Verteilungen:

$$X \sim E_d(\mu, \Sigma, \psi) \Leftrightarrow A^{-1}(X - \mu) \sim S_d(\psi),\ \ A \in \mathbb{R}^{d\times d}, AA^T = \Sigma$$

Theorem 4 *(Stochastische Darstellung der elliptischen Verteilung)*

Sei $X \in \mathbb{R}^d$ ein d-dimensinaler Zufallsvektor.
$X \sim E_d(\mu, \Sigma, \psi)$ dann und nur dann wenn $X \stackrel{d}{=} \mu + RAS$, wobei $S \in \mathbb{R}^k$ ist ein auf der Einheitskugel $\mathcal{S}^{k-1}$ gleichverteilter Zufallsvektor, $R \geq 0$ ist eine von S unahängige nicht negative Zufallsvariable, $A \in \mathbb{R}^{d \times k}$ ist eine konstante Matrix ($\Sigma = AA^T$) und $\mu \in \mathbb{R}^d$ ist ein konstanter Vektor.

Simulation einer elliptischen Verteilung:

(i) Simuliere s aus einer gleichmäßig verteilten Zufallsvektor in $\mathcal{S}^{d-1}$ (zB. in dem y aus einer multivariaten Standard Normalverteilung $Y \sim N_d(0, I)$ simuliert und $s = y/||y||$ gesetzt wird).

(ii) Simuliere r aus R.

(iii) Setze $x = \mu + rAs$.

Beispiel 6 *(Multivariate Normalverteilung)*

Sei $X \sim N(\mu, \Sigma)$. *Es existiert eine Matrix* $A \in \mathbb{R}^{d \times k}$, *sodass* $X \stackrel{d}{=} \mu + AZ$ *wobei* $Z \in N_k(0, I)$ *und* $AA^T = \Sigma$. *Weiters gilt* $Z = RS$ *wobei* S *ein gleichmäßig verteilter Zufallsvektor in* $\mathcal{S}^{k-1}$ *ist und* $R^2 \sim \chi^2_k$. *Daraus folgt* $X \stackrel{d}{=} \mu + RAS$ *und daher* $X \sim E_d(\mu, \Sigma, \psi)$ *mit* $\psi(x) = \exp\{-x/2\}$.

Beispiel 7 *(Multivariate normal variance mixture)*

Sei $Z \sim N_d(0, I)$ *ein normal-verteilter Zufallsvektor.* Z *ist sphärisch-verteilt mit stochastischer Darstellung* $Z \stackrel{d}{=} VS$ *wobei* $V^2 = ||Z||^2 \sim \chi^2_d$. *Sei* $X = \mu + WAZ$ *eine Varianz-gemischte Normalverteilung. Dann gilt* $X \stackrel{d}{=} \mu + VWAS$ *wobei* $V^2 \sim \chi^2_d$ *und* VW *eine nicht-negative von* S *unabhängige ZV ist. D.h.,* X *ist elliptisch verteilt mit* $R = VW$.

Theorem 5 *(Eigenschaften der elliptischen Verteilung)*

Sei $X \sim E_k(\mu, \Sigma, \psi)$. X hat folgende Eigenschaften:

Lineare Kombinationen:

Für $B \in \mathbb{R}^{k \times d}$ und $b \in \mathbb{R}^k$ gilt:

$$BX + b \in E_k(B\mu + b, B\Sigma B^T, \psi).$$

Randverteilungen:

Setze $X^T = \left(X^{(1)^T}, X^{(2)^T}\right)$ für $X^{(1)^T} = (X_1, X_2, \ldots, X_n)^T$ und $X^{(2)^T} = (X_{n+1}, X_{n+2}, \ldots, X_k)^T$ und analog $\mu^T = \left(\mu^{(1)^T}, \mu^{(2)^T}\right)$ sowie $\Sigma = \begin{pmatrix} \Sigma^{(1,1)} & \Sigma^{(1,2)} \\ \Sigma^{(2,1)} & \Sigma^{(2,2)} \end{pmatrix}$. Es gilt dann $X_1 \sim N_n\left(\mu^{(1)}, \Sigma^{(1,1)}, \psi\right)$ und $X_2 \sim N_{k-n}\left(\mu^{(2)}, \Sigma^{(2,2)}, \psi\right)$.

Bedingte Verteilungen:

Wenn Σ regulär, dann ist auch die bedingte Verteilung $X^{(2)}\Big|X^{(1)} = x^{(1)}$ elliptisch verteilt:

$$X^{(2)}\Big|X^{(1)} = x^{(1)} \sim N_{k-n}\Big(\mu^{(2,1)}, \Sigma^{(22,1)}, \tilde{\psi}\Big)$$

wobei

$$\mu^{(2,1)} = \mu^{(2)} + \Sigma^{(2,1)}\Big(\Sigma^{(1,1)}\Big)^{-1}\Big(x^{(1)} - \mu^{(1)}\Big)$$

und

$$\Sigma^{(22,1)} = \Sigma^{(2,2)} - \Sigma^{(2,1)}\Big(\Sigma^{(1,1)}\Big)^{-1}\Sigma^{(1,2)}.$$

Typischerwise sind $\tilde{\psi}$ und ψ unterschiedlich
(siehe Fang, Katz und Ng 1987).

Quadratische Formen:

Wenn Σ regulär, dann gilt

$$D^2 = (X - \mu)^T \Sigma^{-1} (X - \mu) \sim R^2.$$

wobei R die nicht-negative ZV aus der stochastischen Darstellung $Y = RS$ der spherischen Verteilung Y mit $S \sim U\left(\mathcal{S}^{(d-1)}\right)$ und $X = \mu + AY$ ist. Die Zufallsvariable D heißt *Mahalanobis Distanz*.

Faltung:

Seien $X \sim E_k(\mu, \Sigma, \psi)$ und $Y \sim E_k(\tilde{\mu}, \Sigma, \tilde{\psi})$ zwei unabhängige Zufallsvektoren. Es gilt dann $X + Y \sim E_k(\mu + \tilde{\mu}, \Sigma, \bar{\psi})$ wobei $\bar{\psi} = \psi\tilde{\psi}$.

Achtung: Σ muss i.a. dieselbe für X und Y sein.

Anmerkung: Aus $X \sim E_k(\mu, I_k, \psi)$ folgt nicht, dass die Komponenten von X unabhängig sind. Die Komponenten von X sind dann und nur dann unabhängig wenn X multivariat normalverteilt mit der Einheitsmatrix als Kovarianzmatrix ist.