1 Aufgabe T-1

Gegeben seien zwei normalverteilte Zufallsvariablen $X_1 \sim N(\mu_1, \sigma_1^2)$ und $X_2 \sim N(\mu_2, \sigma_2^2)$ mit pdf $p_1(x)$ bzw. $p_2(x)$. Bestimmen Sie x^* (als Funktion der μ_i, σ_i, sodass

$$\int_{-\infty}^{x^*} p_1(x')dx' = \int_{x^*}^{+\infty} p_2(x')dx',$$

d.h., $Pr\{X_1 \leq x^*\} = Pr\{X_2 \geq x^*\} = 1 - Pr\{X_2 \leq x^*\}$.

Hinweis: es ist nicht erforderlich, zu integrieren!

1.1 Lösung

Bezeichne F_1 die Verteilungsfunktion (cdf) von X_1 und F_2 die Verteilungsfunktion von X_2. Die Werte dieser Verteilungsfunktionen lassen sich durch Rückführung auf die Standardnormalverteilung berechnen:

$$F_i(x) = G(\frac{x - \mu_i}{\sigma_i}),$$

wobei G die Verteilungsfunktion der Standardnormalverteilung bezeichnet. Die Aufgabe lässt sich also folgendermaßen formulieren

$$G(\frac{x^* - \mu_1}{\sigma_1}) = 1 - G(\frac{x^* - \mu_2}{\sigma_2})$$

Zusammen mit $G(-x) = 1 - G(x)$ erhält man

$$G(\frac{x^* - \mu_1}{\sigma_1}) = G(-\frac{x^* - \mu_2}{\sigma_2}),$$

und folglich

$$\frac{x^* - \mu_1}{\sigma_1} = -\frac{x^* - \mu_2}{\sigma_2}.$$

Nach einigen einfachen Umformungen erhält man schließlich das gewünschte Ergebnis

$$x^* = \frac{\mu_2\sigma_1 + \mu_1\sigma_2}{\sigma_1 + \sigma_2}.$$

2 Aufgabe T-2

Zeigen Sie, dass für das zentrale zweite Moment der Zufallsvariablen X (Eq. 30) gilt

$$Var(X) = \mathcal{E}[(X - \mathcal{E}[X])^2] = \mathcal{E}[X^2] - \mathcal{E}[X]^2$$

2.1 Lösung

$$\mathcal{E}[(X - \mathcal{E}[X])^2] = \mathcal{E}[X^2 - 2X\mathcal{E}[X] + \mathcal{E}[X]^2].$$

Beachtet man, dass der Ausdruck $\mathcal{E}[X] = \mu$ eine Konstante darstellt, und dass die Erwartung einer Konstanten die Konstante selbst ist, d.h. $\mathcal{E}[\alpha] = \alpha$ (siehe Eq. 28), so ergibt sich mit der Additivität des Erwartungsoperators (Eq. 32)

$$\mathcal{E}[X^2 - 2X\mu + \mu^2] = \mathcal{E}[X^2] - 2\mu\mathcal{E}[X] + \mu^2 = \mathcal{E}[X^2] - 2\mu^2 + \mu^2.$$

3 Aufgabe T-3

Leiten Sie Eq. 34 (Varianz einer Summe von Zufallsvariablen)und Eq. 35 (Kovarianz) her.

3.1 Lösung

Bezeichne im folgenden $\mu_x = \mathcal{E}[X]$ und $\mu_y = \mathcal{E}[Y]$.

$$\begin{aligned} \mathcal{E}[(X + Y - \mathcal{E}[X + Y])^2] &= \\ \mathcal{E}[(X - \mu_x + Y - \mu_y)^2] &= \\ \mathcal{E}[(X - \mu_x)^2 + 2(X - \mu_x)(Y - \mu_y) + (Y - \mu_y)^2] &= \\ \sigma_x^2 + \sigma_y^2 + 2\mathcal{E}[(X - \mu_x)(Y - \mu_y)]. \end{aligned}$$

Der letzte Term, die Kovarianz, läßt sich nun analog zu Aufgabe 2 vereinfachen.

4 Aufgabe T-4

Zeigen Sie, dass der Schätzer der Populatinsvarianz Eq. 41 erwartungstreu ist.

4.1 Lösung

$$\begin{aligned}
\mathcal{E}[\hat{\sigma}_X^2] &= \mathcal{E}[\frac{1}{N}\sum_{i=1}^{N}(X_i-\mu)^2] = \\
\frac{1}{N}\sum_{i=1}^{N}\mathcal{E}[(X_i-\mu)^2] &= \frac{N}{N}\sigma^2.
\end{aligned}$$

5 Aufgabe T-5

Zeigen Sie, dass der Schätzer der Populatinsvarianz Eq. 42 erwartungstreu ist.

5.1 Lösung

Der Beweis basiert auf der Beobachtung, dass X_i, X_j im Falle $i \neq j$ unabhängig, jedoch für $i = j$ abhängig sind (da im letzteren Fall beide für jede mögliche Realisation denselben Wert annehmen müssen).

$$\begin{aligned}
\mathcal{E}[\sum_{i=1}^{N}(X_i-\hat{\mu})^2] = \sum_{i=1}^{N}(\mathcal{E}[X_i^2] - 2\mathcal{E}[X_i\hat{\mu}] + \mathcal{E}[\hat{\mu}^2]) &= \\
N\mathcal{E}[X^2] - 2\mathcal{E}[X_i\sum_{j=1}^{N}X_j] + \frac{1}{N}\mathcal{E}[\sum_{k=1}^{N}\sum_{l=1}^{N}X_kX_l] &= \\
N\mathcal{E}[X^2] - 2(N-1)\mathcal{E}[X]^2 - 2\mathcal{E}[X^2] + (N-1)\mathcal{E}[X]^2 + \mathcal{E}[X^2] &= \\
N\mathcal{E}[X^2] - (N-1)\mathcal{E}[X]^2 - \mathcal{E}[X^2] &= \\
(N-1)(\mathcal{E}[X^2] - \mathcal{E}[X]^2).
\end{aligned}$$

6 Aufgabe T-6

Leiten Sie das Bayes Theorem in der Form Eq. 56 her.

6.1 Lösung

Aus den bedingten Wahrscheinlichkeiten

$$\begin{aligned} P(A|B) &= \frac{P(A,B)}{P(B)} \\ P(B|A) &= \frac{P(A,B)}{P(A)} \end{aligned}$$

kann die *joint probability* $P(A,B)$ wie folgt berechnet werden

$$P(A,B) = P(A|B)P(B)$$
$$P(A,B) = P(B|A)P(A).$$

Durch Gleichsetzen der rechten Seiten und Division durch einen der *priors* erhält man das Bayes Theorem.

7 Aufgabe T-7

Beweisen Sie Eq. 60 und Eq. 59 (in dieser Reihenfolge).

7.1 Lösung

Der Nenner in

$$P(\omega_j|X_i) = \frac{P(X_i|\omega_j)P(\omega_j)}{P(X_i)}$$

(Bayes rule) ergibt sich als Randwahrscheinlichkeit von X_i bzg. ω, d.h. als

$$P(X_i) = \sum_{j=1}^{c} P(X_i, \omega_j).$$

Andererseits gilt

$$P(X_i, \omega_j) = P(X_i|\omega_j)P(\omega_j),$$

und somit

$$P(X_i) = \sum_{j=1}^{c} P(X_i|\omega_j)P(\omega_j).$$

Eq. 59 folgt nun unmittlebar aus Eq. 60 und der Bayes rule

$$\sum_{j=1}^{c} P(\omega_j|X_i) = \frac{\sum_{j=1}^{c} P(X_i|\omega_j)P(\omega_j)}{\sum_{k=1}^{c} P(X_i|\omega_k)P(\omega_k)} = 1.$$

8 Aufgabe T-8

Zeigen Sie, dass die *evidence* $p(x)$ in Eq. 62 die auf Seite 44 oben angegebenen Eigenschaften erfüllt und daher eine Dichtefunktion (*pdf*) ist.

8.1 Lösung (Skizze)

Die *evidence*

$$p(x) = \sum_{j=1}^{c} p(x|\omega_j)P(\omega_j)$$

ist definiert als mit den korrespondierenen *priors* gewichtete Summe der *class conditional pdfs*. Der Beweis dafür, dass die *evidence* ebenfalls eine pdf ist, beruht darauf, dass

- alle $p(x|\omega_j)$ die Eigenschaften einer pdf erfüllen
- die Gewichtungsfaktoren $P(\omega_j)$ nicht-negativ sind und zu 1 summieren,
- das Integral und die Ableitung lineare Operatoren sind.

Im besonderen gilt

$$p(x) \geq 0 \quad \forall(x).$$

Die Verteilungsfunktion der *evidence* berechnet wie folgt

$$F(x) = \int_{-\infty}^{x} \sum_{j=1}^{c} P(\omega_j)p(x'|\omega_j)dx' =$$
$$\sum_{j=1}^{c} P(\omega_j) \int_{-\infty}^{x} p(x'|\omega_j)dx' =$$
$$\sum_{j=1}^{c} P(\omega_j)F(x|\omega_j),$$

d.h. als mit den korrespondierenen *priors* gewichtete Summe der *class conditional cdfs*. Daher gilt

$$F(+\infty) = \sum_{j=1}^{c} P(\omega_j)F(+\infty|\omega_j) = \sum_{j=1}^{c} P(\omega_j) = 1.$$

Die Beweise für die restlichen Eigenschaften einer pdf erhält man analog.

9 Aufgabe T-9

Seien X, Y, Z drei Zufallsvariablen, und bezeichne weiters A, B, C beliebige Ereignisse bzg. dieser Zufallsvariablen (z.B. $X = i, Y = j, Z = k$, analog zu Eq. 52).

Die Zufallsvariablen X, Y sind bedingt unabhängig (*conditionally independent*) bezüglich Z, wenn für beliebige Ereignisse A, B, C gilt, dass

$$P(A, B|C) = P(A|C)P(B|C)$$

(“,” bindet stärker als “|”).

Eine andere Definition der bedingten Unabhängigkeit von X und Y ist durch

$$P(A|B, C) = P(A|C)$$

gegeben. Zeigen Sie, dass diese beiden Definitionen äquivalent sind.

9.1 Lösung

Def2 $\Rightarrow$ Def1

$$P(A, B|C) = \frac{P(A, B, C)}{P(C)} \frac{P(B, C)}{P(B, C)} = P(A|B, C)P(B|C) = P(A|C)P(B|C).$$

Def1 $\Rightarrow$ Def2

$$P(A|B, C) = \frac{P(A, B, C)}{P(B, C)} \frac{P(C)}{P(C)} = \frac{P(A, B|C}{P(B|C)} = \frac{P(A|C)P(B|C)}{P(B|C)} = P(A|C).$$

10 Aufgabe T-10

Drücken Sie die empirische Kovarianzmatrix $\hat{\mathbf{C}}$ durch die empirische Autokorrelations-Matrix $\hat{\mathbf{S}} = \frac{1}{N}\mathbf{X}\mathbf{X}^T$ und die Matrix $\hat{\mathbf{m}}\hat{\mathbf{m}}^T$ aus ($\hat{\mathbf{m}}$...empirisches Mittel).

10.1 Lösung

Bezeichne $\mathbf{1}_N = (1, \ldots, 1)^T$ den Spaltenvektor der Länge N mit allen Einträgen gleich 1. Es gilt

$$\tilde{\mathbf{X}} = (\mathbf{X} - \hat{\mathbf{m}}, \ldots, \hat{\mathbf{m}}) = \mathbf{X} - \hat{\mathbf{m}}\mathbf{1}_N^T.$$

Somit erhalten wir

$$\begin{aligned}\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T &= (\mathbf{X} - \hat{\mathbf{m}}\mathbf{1}_N^T)(\mathbf{X} - \hat{\mathbf{m}}\mathbf{1}_N^T)^T \\ &= \mathbf{X}\mathbf{X}^T - \hat{\mathbf{m}}\mathbf{1}_N^T\mathbf{X}^T - \mathbf{X}\mathbf{1}_N\hat{\mathbf{m}}^T + \hat{\mathbf{m}}\mathbf{1}_N^T\mathbf{1}_N\hat{\mathbf{m}}^T,\end{aligned}$$

und, da $\mathbf{X}\mathbf{1}_N = N\hat{\mathbf{m}}$ und $\mathbf{1}_N^T\mathbf{1}_N = N$,

$$\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T = \mathbf{X}\mathbf{X}^T - N\mathbf{m}\mathbf{m}^T.$$

Division durch $N-1$ liefert schließlich

$$\begin{aligned}\hat{\mathbf{C}} &= \frac{1}{N-1}\tilde{\mathbf{X}}\tilde{\mathbf{X}}^T \\ &= \frac{N}{N-1}(\hat{\mathbf{S}} - \mathbf{m}\mathbf{m}^T).\end{aligned}$$

11 Aufgabe T-11

Zeigen Sie, dass die Komponenten $X_i, 1 \leq i \leq p$ des multivariat normalverteilten Zufallsvektors $\vec{X} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ unabhängig sind, falls $\boldsymbol{\Sigma}$ eine Diagonalmatrix ist.

Berechnen Sie dazu die *marginal pdfs* $p(x_i)$ und zeigen sie, dass die *joint pdf* $p(\mathbf{x})$ in diese faktorisiert, d.h. $p(\mathbf{x}) = \prod_{i=1}^{p} p(x_i)$.

11.1 Lösung

Im Fall einer diagonalen Kovarianzmatrix $\boldsymbol{\Lambda} = diag(\lambda_1, \ldots, \lambda_p)$ läßt sich die joint pdf

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{p\frac{1}{2}} |\boldsymbol{\Lambda}|^{\frac{1}{2}}} \exp(-\frac{1}{2}(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Lambda}^{-1} (\mathbf{x} - \boldsymbol{\mu})) \tag{1}$$

äquivalent als

$$p(\mathbf{x}) = \frac{1}{(2\pi)^{p\frac{1}{2}} (\prod_{i=1}^{p} \lambda_i)^{\frac{1}{2}}} \exp(-\frac{1}{2} \sum_{i=1}^{p} \frac{(x_i - \mu_i)^2}{\lambda_i}) \tag{2}$$

formulieren.

Eq. 2 läßt sich, wie man leicht sieht, als Produkt von p Faktoren darstellen

$$p(\mathbf{x}) = \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi\lambda_i}} \exp(-\frac{1}{2} \frac{(x_i - \mu_i)^2}{\lambda_i}), \tag{3}$$

webei der j-te Faktor der pdf einer univariaten Normalverteilung mit Mittel μ_i und Varianz $\sigma_i^2 = \lambda_i$ entspricht.

Die Randverteilung der j-ten Komponente $p(x_j)$ erhält man durch Integration über alle anderen Komponenten

$$\begin{aligned} p(x_j) &= \int \ldots \int \prod_{i=1}^{p} \frac{1}{\sqrt{2\pi\lambda_i}} \exp(-\frac{1}{2} \frac{(x_i - \mu_i)^2}{\lambda_i}) dx_1 \ldots dx_{j-1} dx_{j+1} \ldots dx_p \\ &= \frac{1}{\sqrt{2\pi\lambda_j}} \exp(-\frac{1}{2} \frac{(x_j - \mu_j)^2}{\lambda_j}) \prod_{i \neq j} \int \frac{1}{\sqrt{2\pi\lambda_i}} \exp(-\frac{1}{2} \frac{(x_i - \mu_i)^2}{\lambda_i}) dx_i. \end{aligned} \tag{4}$$

Der letzte Schritt folgt daraus, dass der i-te Faktor des Produkts ausschließlich von der i-ten Varianblen x_i abhängt. Da die Integranden pdfs von Normalverteilungen sind, müssen alle Integrale den Wert 1 annehmen. Somit haben wir

$$p(x_j) = \frac{1}{\sqrt{2\pi\lambda_j}} \exp(-\frac{1}{2} \frac{(x_j - \mu_j)^2}{\lambda_j}), \tag{5}$$

d.h. die Faktoren in Eq. 3 sind durch die marginal densities $p(x_j)$ gegeben:

$$p(\mathbf{x}) = \prod_{i=1}^{p} p(x_j). \tag{6}$$