

Ue Statistik u. Wahrscheinlichkeitsth. f. Inf. 107.251 W 2005 6 http://www.statistik.tuwien.ac.at/RISueb/	Di 12:00 – 18:00
	8.Blatt
Werner GURKER Tel.: 58801-10724 Spr.: Di u. Do 11-12 e-mail: W.Gurker@tuwien.ac.at	13. Dez. 2005

1. (a) Zeigen Sie für eine nichtnegative stochastische Größe mit $\mu = \mathbb{E}(X)$ die *Markoff'sche Ungleichung*:

$$W\{X \geq a\} \leq \frac{\mu}{a}, \quad a > 0$$

Hinweis: Vgl. Sie den Beweis der Tschebyscheff'schen Ungleichung.

- *(b) Wie folgt aus der Markoff'schen die Tschebyscheff'sche Ungleichung?
2. (a) Vergleichen Sie (rechnerisch/graphisch) für eine nach $N(\mu, \sigma^2)$ verteilte sG X die Wahrscheinlichkeiten:

$$Q(x) := W\{|X - \mu| \geq x\sigma\}, \quad x \geq 0$$

mit den aus der Tschebyscheff'schen Ungleichung resultierenden Werten.

- *(b) Wiederholen Sie (a) für eine nach Ex_τ verteilte sG.
3. Die gemeinsame Dichte von X und Y sei gegeben wie folgt:

$$f(x, y) = \begin{cases} x + y & 0 < x < 1, \ 0 < y < 1 \\ 0 & \text{sonst} \end{cases}$$

- (a) Bestimmen Sie die Randdichten von X und Y .
- (b) Sind X und Y (stochastisch) unabhängig?
- (c) Berechnen Sie $W\{X + Y \leq 1\}$.
4. Fortsetzung von **Bsp 3:**
- (a) Berechnen Sie $\text{Cov}(X, Y)$.
- (b) Wie groß ist die Korrelation ρ_{XY} ?
- *(c) Ermitteln Sie die Regressionsfunktionen von Y bezüglich X und von X bezüglich Y und stellen Sie beide graphisch dar.
5. X und Y seien bivariat normalverteilt mit den Parametern $\mu_1 = 3$, $\mu_2 = 1$, $\sigma_1^2 = 16$, $\sigma_2^2 = 25$ und $\rho = \frac{3}{5}$. Bestimmen Sie die folgenden Wahrscheinlichkeiten:
- (a) $W\{3 < Y < 8\}$
- (b) $W\{3 < Y < 8 | X = 7\}$
- (c) $W\{-3 < X < 3\}$
- (d) $W\{-3 < X < 3 | Y = -4\}$
6. X und Y seien bivariat normalverteilt mit den Parametern $\mu_1 = 5$, $\mu_2 = 10$, $\sigma_1^2 = 1$, $\sigma_2^2 = 25$ und $\rho > 0$. Wenn $W\{4 < Y < 16 | X = 5\} = 0.954$, wie groß ist dann ρ ?

Beispiel(teil)e mit (*) werden im Konversatorium behandelt.

Lösungen zum 8. Blatt

1. (a) Beweis nur für eine stetige sG mit Dichte f (im diskreten oder gemischten Fall argumentiert man analog):

$$\begin{aligned}\mathbb{E}(X) &= \int_0^{\infty} x f(x) dx = \underbrace{\int_{0 < x < a} x f(x) dx}_{\geq 0} + \int_{x \geq a} x f(x) dx \\ &\geq \int_{x \geq a} x f(x) dx \geq a \int_{x \geq a} f(x) dx = a W\{X \geq a\}\end{aligned}$$

- *(b) Ist X eine (beliebige) stochastische Größe mit $\mathbb{E}(X) = \mu$ und $\text{Var}(X) = \sigma^2$, so gilt nach der Markoff'schen Ungleichung ($\varepsilon > 0$):

$$W\{|X - \mu| \geq \varepsilon\} = W\{(X - \mu)^2 \geq \varepsilon^2\} \leq \frac{\mathbb{E}[(X - \mu)^2]}{\varepsilon^2} = \frac{\text{Var}(X)}{\varepsilon^2}$$

Dies ist aber die Tschebyscheff'sche Ungleichung.

2. Allgemein gilt für $X \sim F$ und $x \geq 0$:

$$\begin{aligned}W\{|X - \mu| \geq x\sigma\} &= 1 - W\{|X - \mu| < x\sigma\} \\ &= 1 - W\{\mu - x\sigma < X < \mu + x\sigma\} \\ &= 1 - [F((\mu + x\sigma) -) - F(\mu - x\sigma)]\end{aligned}$$

Bem.: $F(x-)$ ist der linksseitige Limes von F an der Stelle x ; für stetiges X ist $F(x-) = F(x)$.

- (a) $X \sim N(\mu, \sigma^2)$: $\mathbb{E}(X) = \mu$, $\text{Var}(X) = \sigma^2$.

$$\begin{aligned}W\{|X - \mu| \geq x\sigma\} &= 1 - [F(\mu + x\sigma) - F(\mu - x\sigma)] \\ &= 1 - [\Phi(x) - \Phi(-x)] \\ &= 1 - [2\Phi(x) - 1] \\ &= 2[1 - \Phi(x)], \quad x \geq 0\end{aligned}$$

Die Tschebyscheff-Schranke ist $1/x^2$; sie ist nur für $x > 1$ informativ.

x	$Q(x)$	$1/x^2$
0.0	1.0000	∞
0.5	0.6171	4.0000
1.0	0.3173	1.0000
1.5	0.1336	0.4444
2.0	0.0455	0.2500
2.5	0.0124	0.1600
3.0	0.0027	0.1111
3.5	0.0005	0.0816
4.0	0.0001	0.0625

* (b) $X \sim Ex_\tau: \mathbb{E}(X) = \tau, \text{Var}(X) = \tau^2.$

$$W\{|X - \tau| \geq x\tau\} = 1 - [F(\tau(1+x)) - F(\tau(1-x))]$$

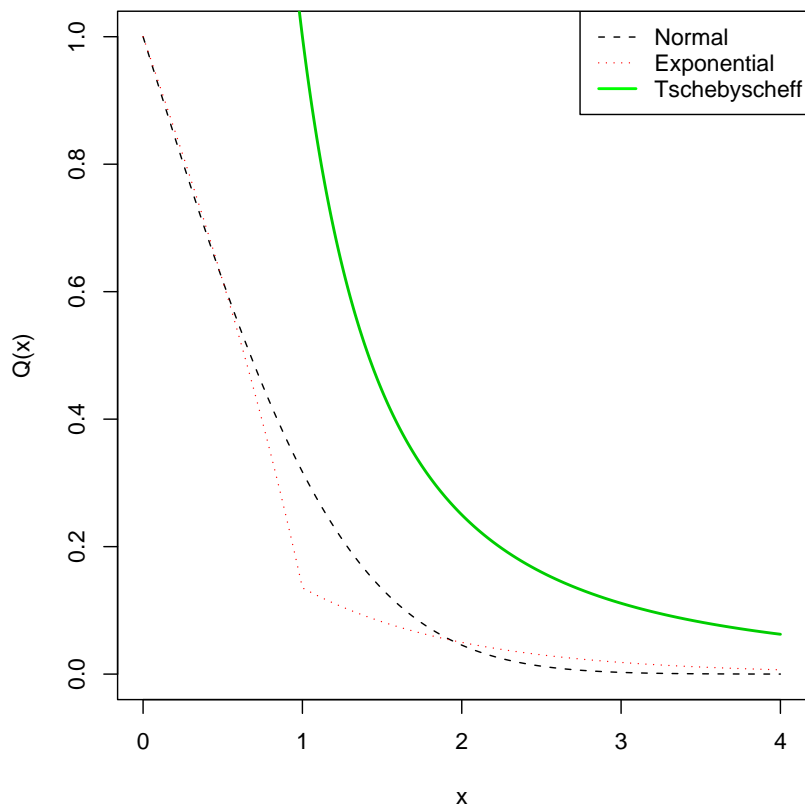
Für die Verteilungsfunktion einer Ex_τ -Verteilung gilt:

$$F(x) = (1 - e^{-x/\tau})I_{[0,\infty)}(x)$$

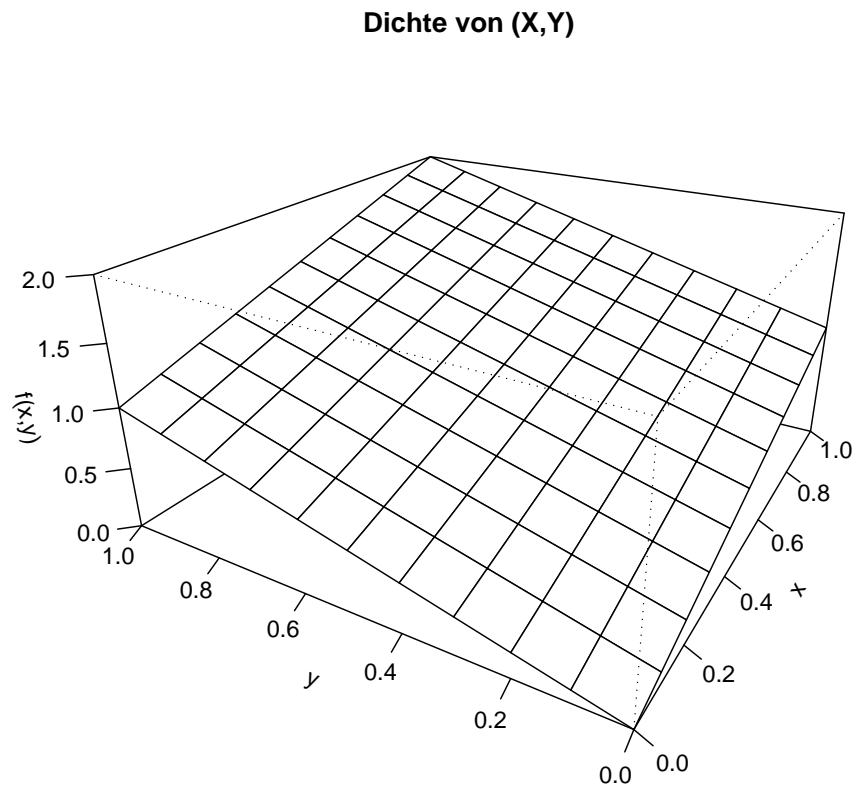
Damit folgt:

$$W\{|X - \tau| \geq x\tau\} = \begin{cases} 1 + e^{-(1+x)} - e^{-(1-x)} & 0 \leq x < 1 \\ e^{-(1+x)} & x \geq 1 \end{cases}$$

x	$Q(x)$	$1/x^2$
0.0	1.0000	∞
0.5	0.6166	4.0000
1.0	0.1353	1.0000
1.5	0.0821	0.4444
2.0	0.0498	0.2500
2.5	0.0302	0.1600
3.0	0.0183	0.1111
3.5	0.0111	0.0816
4.0	0.0067	0.0625



3.



(a) Die Randdichte von X bestimmt man durch Integration nach y :

$$f_1(x) = \int_0^1 (x+y) dy = x + \frac{1}{2}, \quad 0 < x < 1$$

Analog bestimmt man die Randdichte von Y :

$$f_2(y) = \int_0^1 (x+y) dx = y + \frac{1}{2}, \quad 0 < y < 1$$

(b) Wegen $f(x,y) \neq f_1(x)f_2(y)$ sind X und Y nicht unabhängig.

(c)

$$\begin{aligned} W\{X+Y \leq 1\} &= \int_0^1 \int_0^{1-x} (x+y) dy dx = \int_0^1 \left[x(1-x) + \frac{(1-x)^2}{2} \right] dx \\ &= \int_0^1 \left(\frac{1}{2} - \frac{x^2}{2} \right) dx = \frac{1}{3} \end{aligned}$$

4. (a) Zunächst bestimmt man den Erwartungswert von X :

$$\mathbb{E}(X) = \int_0^1 x \left(x + \frac{1}{2} \right) dx = \frac{7}{12}$$

Aus Symmetriegründen gilt $\mathbb{E}(Y) = \mathbb{E}(X)$.

$$\begin{aligned} \mathbb{E}(XY) &= \int_0^1 \int_0^1 xy(x+y) dx dy = \int_0^1 \int_0^1 (x^2y + xy^2) dx dy \\ &= \int_0^1 \left(\frac{y}{3} + \frac{y^2}{2} \right) dy = \frac{1}{3} \end{aligned}$$

Mit dem Verschiebungssatz für Kovarianzen folgt nun:

$$\text{Cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y) = \frac{1}{3} - \frac{49}{144} = -\frac{1}{144}$$

- (b) Für die Varianz von X (= Varianz von Y) gilt:

$$\mathbb{E}(X^2) = \int_0^1 x^2 \left(x + \frac{1}{2} \right) dx = \frac{5}{12} \quad \implies \quad \text{Var}(X) = \frac{5}{12} - \frac{49}{144} = \frac{11}{144}$$

Damit folgt:

$$\rho_{XY} = \frac{\text{Cov}(X, Y)}{\sqrt{\text{Var}(X) \text{Var}(Y)}} = \frac{-1/144}{\sqrt{11/144 \cdot 11/144}} = -\frac{1}{11}$$

- *(c) Die Dichte von Y bedingt durch $X = x$ ist gegeben wie folgt:

$$f(y|x) = \frac{f(x, y)}{f_1(x)} = \frac{x+y}{x+1/2}, \quad 0 < y < 1$$

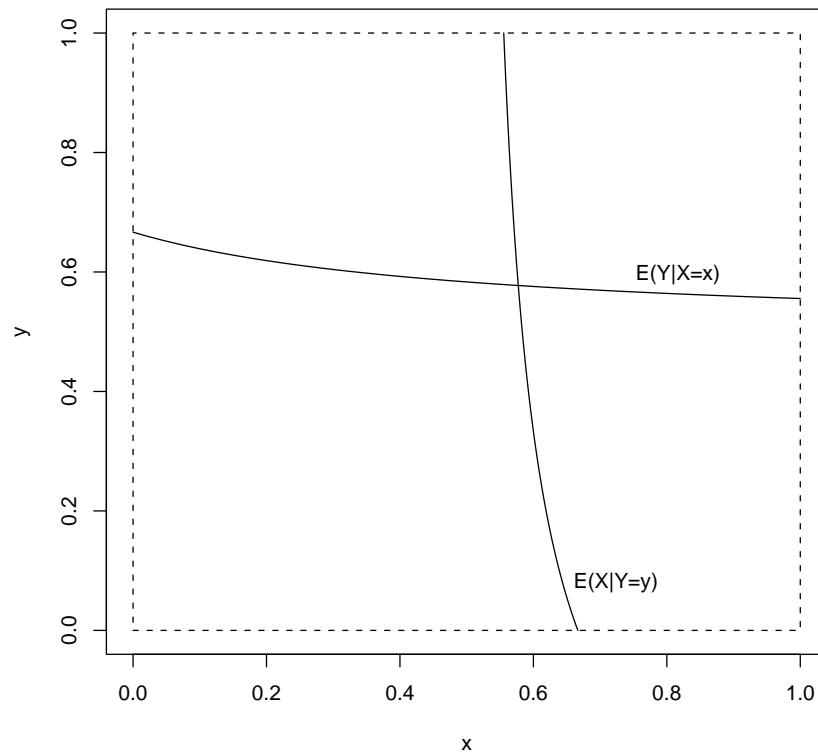
Die Regressionsfunktion von Y bezüglich X ist der bedingte Erwartungswert:

$$\mathbb{E}(Y|X = x) = \int_0^1 y f(y|x) dy = \int_0^1 \frac{y(x+y)}{x+1/2} dy = \frac{3x+2}{6x+3}, \quad 0 < x < 1$$

Analog bestimmt man die Regressionsfunktion von X bezüglich Y :

$$\mathbb{E}(X|Y = y) = \frac{3y+2}{6y+3}, \quad 0 < y < 1$$

Regressionsfunktionen



5. (a) Für die Randverteilung von Y gilt $Y \sim N(1, 25)$; damit folgt:

$$W\{3 < Y < 8\} = \Phi\left(\frac{8-1}{5}\right) - \Phi\left(\frac{3-1}{5}\right) = 0.2638$$

- (b) Für die bedingte Verteilung von $Y|X = 7$ gilt:

$$Y|X = 7 \sim N\left(1 + \frac{3}{5} \cdot \frac{5}{4} \cdot (7-3), \left(1 - \frac{9}{25}\right) \cdot 25\right) = N(4, 16)$$

Damit folgt:

$$W\{3 < Y < 8|X = 7\} = \Phi\left(\frac{8-4}{4}\right) - \Phi\left(\frac{3-4}{4}\right) = 0.4401$$

- (c) Für die Randverteilung von X gilt $X \sim N(3, 16)$; damit folgt:

$$W\{-3 < X < 3\} = \Phi\left(\frac{3-3}{4}\right) - \Phi\left(\frac{-3-3}{4}\right) = 0.4332$$

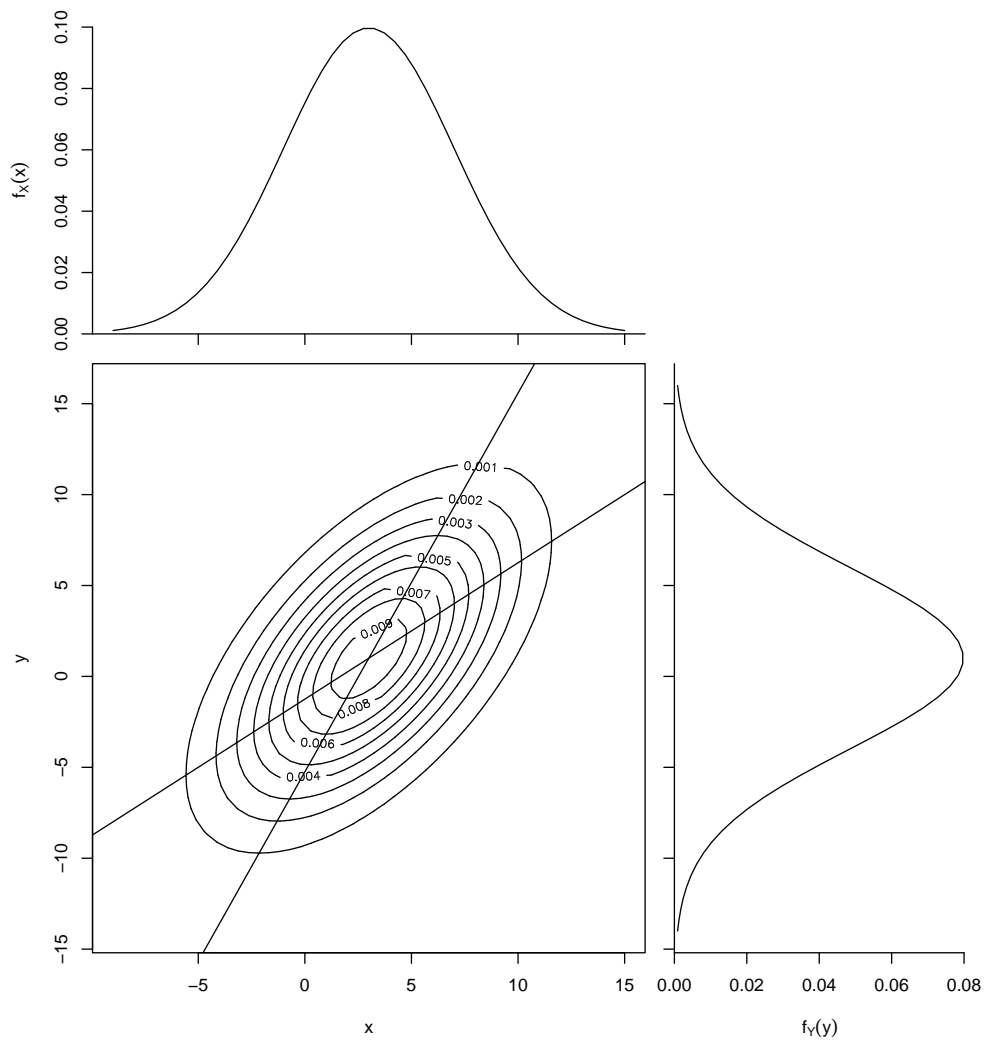
- (d) Für die bedingte Verteilung von $X|Y = -4$ gilt:

$$X|Y = -4 \sim N\left(3 + \frac{3}{5} \cdot \frac{4}{5} \cdot (-4-1), \left(1 - \frac{9}{25}\right) \cdot 16\right) = N\left(\frac{3}{5}, \left(\frac{16}{5}\right)^2\right)$$

Damit folgt:

$$W\{-3 < X < 3|Y = -4\} = \Phi\left(\frac{3-3/5}{16/5}\right) - \Phi\left(\frac{-3-3/5}{16/5}\right) = 0.6431$$

Contourplot von $f(x, y)$ und Randdichten



Die eingezeichneten Geraden sind die Regressionsfunktionen.

6. Laut Angabe gilt:

$$\mathbb{E}(Y|X = 5) = 10 + \rho \frac{5}{1} (5 - 5) = 10; \quad \text{Var}(Y|X = 5) = (1 - \rho^2) \cdot 25$$

$$\Rightarrow W(4 < Y < 16|X = 5) = 2\Phi\left(\frac{6}{5\sqrt{1-\rho^2}}\right) - 1 = 0.954$$

Daraus folgt:

$$\frac{6}{5\sqrt{1-\rho^2}} = u_{(1+0.954)/2} = u_{0.977} \approx 2 \quad \Rightarrow \quad \rho \approx \frac{4}{5}$$