MS-A0509 Grundkurs i sannolikhetskalkyl och statistik
Sammanfattning, del I

G. Gripenberg

Aalto-universitetet

28 januari 2014

Vad är sannolikhet?

- *Relativ frekvens vid upprepningar: Om en fabrik tillverkat* 1 000 000 *exemplar av en produkt av vilka* 5015 *har något fel så är sannolikheten för en felaktighet* 0.005
- *Andelen fall då ett något förekommer: Om i en urna finns* 6 *svarta och* 4 *vita kulor och man slumpmässigt väljer en kula så är sannolikheten att den är svart* $\frac{6}{6+4} = 0.6$.
- *Ett mått på hur troligt man anser något vara: "Sannolikheten för hård vind imorgon är* 70%*."*

Sannolikhet, händelser, utfallsrum

- *Mängden av alla tänkbara resultat av ett "experiment" är* **utfallstummet**, *ofta betecknat med* Ω.
- *Elementen i utfallsrummet, dvs. enskilda resultat av experimentet är* **elementarhändelser**.
- **Händelser** *är delmängder av utfallsrummet.*
- *För varje händelse* $A \subset \Omega$ *finns det en sannolikhet* $\Pr(A)$ *för denna händelse.*
- *Sannolikhetsfunktionen uppfyller följande villkor:*
  - ★ $0 \leq \Pr(A) \leq 1$ *för varje händelse* A.
  - ★ $\Pr(\Omega) = 1$.
  - ★ $\Pr(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} \Pr(A_i)$ *om* $A_j \cap A_k = \emptyset$ *då* $j \neq k$.

Obs

Av antagandena ovan följer bla. också att

- ★ $\Pr(\emptyset) = 0$.
- ★ $\Pr(A \cup B) = \Pr(A) + \Pr(B) - \Pr(A \cap B)$.
- ★ $\Pr(\Omega \setminus A) = 1 - \Pr(A)$.
- ★ $A \subset B \Rightarrow \Pr(A) \leq \Pr(B)$.

Obs!

Då Ω innehåller ändligt många element är det naturligt att att alla delmängder av Ω är händelser men i allmänhet är detta inte alltid möjligt eller ens önskvärt och då är Pr *en funktion definierad i en σ-algebra $\mathcal{A}$ i Ω, dvs en mängd $\mathcal{A}$ med följande egenskaper:*

- $A \in \mathcal{A} \rightarrow A \subset \Omega$,
- $\Omega \in \mathcal{A}$,
- $A \in \mathcal{A} \rightarrow \Omega \setminus A \in \mathcal{A}$,
- $A_i \in \mathcal{A},\ i = 1, 2, \ldots \rightarrow \cup_{i=1}^{\infty} A_i \in \mathcal{A}$.

Oberoende

Händelserna A och B är oberoende ifall

$$\Pr(A \cap B) = \Pr(A) \cdot \Pr(B),$$

och händelserna A_j, $j \in J$ är oberoende om

$$\Pr(A_{j_1} \cap A_{j_2} \cap \ldots \cap A_{j_m}) = \Pr(A_{j_1}) \cdot \Pr(A_{j_2}) \cdot \ldots \cdot \Pr(A_{j_m})$$

alltid då $j_k \in J$, $k = 1, \ldots, m$, $j_p \neq j_q$ då $p \neq q$.

Obs!

Om händelserna A_j, $j \in J$ är oberoende så är A_{j_p} och A_{j_q} oberoende då $j_p \neq j_k$ men om A_{j_p} och A_{j_q} är oberoende för alla $j_p \neq j_k$ så behöver inte händelserna A_j, $j \in J$ vara oberoende.

Betingad sannolikhet

Den betingande sannolikheten för händelsen A givet händelsen B är

$$\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)},$$

då man antar att $\Pr(B) > 0$.
Då händelsen B är given kan man begränsa utfallsrummet från Ω till B och räkna om sannolikheterna för händelserna $A \cap B$ som är delmängder av det nya utfallsrummet.

Total sannolikhet

Om $\cup_{j=1}^{n} A_j = \Omega$, $A_j \cap A_k = \emptyset$ då $j \neq k$ och $\Pr(A_j) > 0$ då $j = 1, \ldots, n$ så gäller

$$\Pr(B) = \sum_{j=1}^{n} \Pr(A_j) \cdot \Pr(B|A_j).$$

Varför? Eftersom $B = B \cap \Omega = \cup_{j=1}^{n} B \cap A_j$ och $(B \cap A_j) \cap (B \cap A_k) = \emptyset$ då $j \neq k$ så är $\Pr(B) = \sum_{j=1}^{n} \Pr(B \cap A_j)$ och enligt defintionen är $\Pr(A_j) \cdot \Pr(B|A_j) = \Pr(B \cap A_j)$.

Bayes formel

Om $\cup_{j=1}^{n} A_j = \Omega$, $A_j \cap A_k = \emptyset$ då $j \neq k$ och $\Pr(A_j) > 0$ då $j = 1, \ldots, n$ så gäller

$$\Pr(A_k|B) = \frac{\Pr(A_k) \cdot \Pr(B|A_k)}{\sum_{j=1}^{n} \Pr(A_j) \cdot \Pr(B|A_j)}.$$

Varför?

$$\Pr(A_k|B) = \frac{\Pr(A_k \cap B)}{\Pr(B)}, \quad \Pr(A_k) \cdot \Pr(B|A_k) = \Pr(A_k \cap B) \quad \textit{och}$$

$$\sum_{j=1}^{n} \Pr(A_j) \cdot \Pr(B|A_j) = \Pr(B).$$

Klassisk sannolikhet och kombinatorik

$$\Pr(A) = \frac{\textit{Antal fall då A inträffar}}{\textit{Totala antalet möjliga fall}}$$

Man antar alltså att varje elementarhändelse är lika sannolik och problemet blir att bestämma hur många element det det finns i utfallsrummet Ω och hur många av dessa hör som till mängden A.

Produktprincipen

Om i en urvalsprocess finns k steg och i steg j finns n_j alternativ, oberoende av vilka val som gjorts i tidigare steg (men vilka alternativen är kan bero på valen) *så är det totala antalet alternativ*

$$n_1 \cdot n_2 \cdot \ldots \cdot n_k.$$

Permutationer, binomialkoefficineter etc.

Om det i en mängd finns n element kan dessa ordnas på

$$n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n,$$

olika sätt. (Kom ihåg: $0! = 1$)
Om man ur en mängd med n element väljer k element och beaktar i vilken ordning elementen väljs, kan detta göras på

$$n \cdot (n-1) \cdot \ldots \cdot (n-k+1) = \frac{n!}{(n-k)!}$$

olika sätt.
Om man ur en en mängd med n element väljer en delmängd med k element, dvs. inte beaktar i vilken ordning elementen väljs, kan detta göras på

$$\binom{n}{k} = \frac{n!}{k!(n-k)!},$$

olika sätt.
I alla dessa fall kan varje element i mängden väljas bara en gång, dvs. om man plockar kulor ur en urna läggs dessa inte tillbaka i urnan.

Upprepning av "experiment"

Antag att ett experiment upprepas n gånger så att händelser vid olika gånger är oberoende. *Då är sannolikheten för att händelsen A inträffar exakt k gånger*

$$\binom{n}{k} \Pr(A)^k \big(1 - \Pr(A)\big)^{n-k}.$$

Slumpvariabler och fördelningsfunktioner

En (reell) **slumpvariabel** *(eller* **stokastisk variabel***) är en funktion $X : \Omega \to \mathbb{R}$ (alltså inte egentligen en variabel) där Ω är ett utfallsrum för ett experiment i vilken en sannolikhet är definierad* och $\{\, \omega \in \Omega : X(\omega) \le x \,\}$ är en händelse för alla $x \in \mathbb{R}$.
Om X är en (reell) slumpvariabel så är dess (kumulativa) fördelningsfunktion funktionen

$$F_X(x) = \Pr(X \le x) = \Pr\big(\{\, \omega \in \Omega : X(\omega) \le x \,\}\big).$$

En funktion $F : \mathbb{R} \to [0,1]$ är en fördelningsfunktion om och endast om

- $0 \le F(x) \le F(y) \le 1$ *då* $x < y$,
- $\lim_{x\to-\infty} F(x) = 0$ *och* $\lim_{x\to\infty} F(x) = 1$,
- $\lim_{y\to x+} F(y) = F(x)$ *då* $x \in \mathbb{R}$.

När F är en fördelningsfunktion för X så gäller dessutom att

- $\lim_{y\to x-} F(y) = \Pr(X < x)$,
- $\lim_{y\to x+} F(y) - \lim_{y\to x-} F(y) = \Pr(X = x)$.

Obs!

Uttryck som $X \le x$ och $X < x$ är formellt sett inte händelser (dvs. delmängder i Ω) men man skriver oftast $\Pr(X \le x)$ istället för det längre uttrycket $\Pr\big(\{\,\omega \in \Omega : X(\omega) \le x\,\}\big)$.

Oberoende slumpvariabler

De (reella) slumpvariablerna X_j, $j \in J$ definierade i samma utfallsrum är oberoende om händelserna $\{\,X_j \le a_j\,\}$, $j \in J$ är oberoende för alla $a_j \in \mathbb{R}$, $j \in J$.

Diskreta slumpvariabler

En (reell) slumpvariabel X är diskret om det finns en mängd $A \subset \mathbb{R}$ och positiva tal $f(a)$, $a \in A$ så att

$$F_X(x) = \sum_{\substack{a \le x\\ a \in A}} f(a).$$

*Detta innebär att $\Pr(X = a) = f(a)$ då $x \in A$ och $\sum_{a \in A} f(a) = 1$ så att $\Pr(X \notin A) = 0$ och mängden A innehåller högst numrerbart många element liksom oftast också utfallsrummet Ω. Funktionen f är **frekvensfunktionen** eller sannolikhetsfunktionen för X.*

Kontinuerliga slumpvariabler

*En slumpvariabel är kontiuerlig om fördelningsfunktionen är kontinuerlig, dvs. om $\Pr(X = a) = 0$ för alla $a \in \mathbb{R}$. Oftast använder man sig ändå av absolut kontinuerliga slumpvariabler för vilka det finns en **täthetsfunktion** f så att*

$$F_X(x) = \int_{-\infty}^{x} f(t)\,\mathrm{d}t.$$

Detta innebär att $f(x) \ge 0$ och $\int_{-\infty}^{\infty} f(t)\,\mathrm{d}t = 1$.

Väntevärde och varians

Om X är en slumpvariabel så är dess väntevärde

$$\mathrm{E}(X) = \sum_x x \Pr(X = x) = \sum_x x f_X(x),$$

då X är en diskret slumpvariable med frekvensfunktion f_X och

$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f_X(x)\,\mathrm{d}x,$$

då X är en absolut kontinerlig slumpvariabel med täthetsfunktion f_X, i båda fallen förutsatt att summan eller integralen existerar (i annat fall kan man skriva $\mathrm{E}(X) = NaN$).
Om g är en mätbar funktion så är

$$\mathrm{E}(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x)\,\mathrm{d}x.$$

*Om slumpvariabeln X har ett väntevärde så är dess **varians***

$$\mathrm{Var}(X) = \mathrm{E}\Big(\big(X - \mathrm{E}(X)\big)^2\Big).$$

Väntevärde och varians, forts.

Om X_1 och X_2 är två slumpvariabler så är

$$\mathrm{E}(c_1X_1 + c_2X_2) = c_1\mathrm{E}(X_1) + c_2\mathrm{E}(X_2),$$

och om X_1 och X_2 är två **oberoende** *slumpvariabler så är*

$$\mathrm{Var}(c_1X_1 + c_2X_2) = c_1^2\mathrm{Var}(X_1) + c_2^2\mathrm{Var}(X_2).$$

Chebshevs olikhet

$$\Pr\left(|X - \mathrm{E}(X)| \geq k\sqrt{\mathrm{Var}(X)}\right) \leq \frac{1}{k^2}, \quad k > 1.$$

Varför?

Låt $g(x) = 1$ om $\frac{|x-\mathrm{E}(X)|}{k\sqrt{\mathrm{Var}(X)}} \geq 1$ och 0 annars. Detta betyder att

$\mathrm{E}(g(X)) = \Pr(|X - \mathrm{E}(X)| \geq k\sqrt{\mathrm{Var}(X)})$. *Nu är $g(x) \leq \left(\frac{|x-\mathrm{E}(X)|}{k\sqrt{\mathrm{Var}(X)}}\right)^2$ så att*

$$\Pr\left(|X - \mathrm{E}(X)| \geq k\sqrt{\mathrm{Var}(X)}\right) = \mathrm{E}(g(X))$$
$$\leq \mathrm{E}\left(\left(\frac{|x - \mathrm{E}(X)|}{k\sqrt{\mathrm{Var}(X)}}\right)^2\right) = \frac{1}{k}^2\frac{\mathrm{Var}(X)}{\mathrm{Var}(X)} = \frac{1}{k^2}.$$

Några viktiga diskreta slumpvariabler och deras fördelningar

- *Jämn diskret fördelning:* $\Pr(X = x_i) = \frac{1}{n}$, $i = 1, 2, \ldots, n$.
- *Bernoullifördelning:* $\Pr(X = 1) = p$, $\Pr(X = 0) = (1 - p)$. *I detta fall är $X(\omega) = 1$ då $\omega \in A \subset \Omega$ och $X(\omega) = 0$ då $\omega \in \Omega \setminus A$ där* $\Pr(A) = p$. *Väntevärdet p och variansen $p(1 - p)$.*
- *Binomialfördelning $X \sim \mathrm{Bin}(n, p)$:* $\Pr(X = k) = \binom{n}{k}p^k(1 - p)^{n-k}$, *$k = 0, 1, \ldots, n$, dvs. X är summan av n oberoende Bernoulli-fördelade slumvariabler.* $\mathrm{E}(X) = np$ *och* $\mathrm{Var}(X) = np(1 - p)$.
- *Poissonfördelning $X \sim \mathrm{Poisson}(\lambda)$:* $\Pr(X = k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$, *$k = 0, 1, 2, \ldots$. Fås som gränsvärde av binomialfördelningen då $np \to \lambda$. Väntevärdet och variansen är λ.*
- *Två varianter av geometrisk fördelning: Ett experiment upprepas tills händelsen A med sannolikheten p inträffar. X_1 är antalet upprepningar och X_2 är antalet upprepningar innan A inträffar så att $X_1 = X_2 + 1$,* $\Pr(X_1 = k) = (1 - p)^{k-1}p$, *$k \geq 1$ och* $\Pr(X_2 = k) = (1 - p)^k p$, *$k \geq 0$.* $\mathrm{E}(X_1) = \frac{1}{p}$, $\mathrm{E}(X_2) = \frac{1-p}{p}$ *och båda har variansen* $\frac{1-p}{p^2}$.

Några viktiga kontinuerliga slumpvariabler och deras fördelningar

- *Likformig kontinuerlig fördelning (där $-\infty < a < b < \infty$):*
$$f_X(x) = \begin{cases} \frac{1}{b-a}, & a < x < b, \\ 0, & x \leq a \quad \text{eller} \quad x >\geq b. \end{cases}$$
Väntevärdet är $\frac{1}{2}(a + b)$ och variansen $\frac{1}{12}(b - a)^2$.
- *Normalfördelning $X \sim N(\mu, \sigma^2)$:*
$$f_x(x) = \frac{1}{\sigma\sqrt{2\pi}}\mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$
Väntevärdet är μ och variansen σ^2.
- *Exponentialfördelning $X \sim \mathrm{Exp}(\lambda)$ med $\lambda > 0$:*
$$f_X(x) = \begin{cases} \lambda\mathrm{e}^{-\lambda x}, & x > 0, \\ 0, & x \leq 0. \end{cases}$$
Väntevärdet är $\frac{1}{\lambda}$ och variansen $\frac{1}{\lambda^2}$.

Centrala gränsvärdessatsen

Ifall slumpvariablerna $X_1, X_2, \ldots$ är oberoende och har samma fördelning så att $\mathrm{E}(X_j) = \mu$ och $\mathrm{Var}(X_j) = \sigma^2$, $j = 1, 2, \ldots$, så gäller

$$\frac{\sum_{j=1}^n X_j - n\mu}{\sigma\sqrt{n}} \sim_a \mathrm{N}(0,1) \quad \text{då } n \to \infty,$$

dvs.

$$\lim_{n\to\infty} \Pr\left(\frac{\sum_{j=1}^n X_j - n\mu}{\sigma\sqrt{n}} \le x\right) = F_{\mathrm{N}(0,1)}(x) \stackrel{\text{def}}{=} \int_{-\infty}^x \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2}\,dt.$$

Normalapproximation

Om X är summan av "tillräckligt" många oberoende slumpvariabler med ändlig varians så är $\dfrac{X - \mathrm{E}(X)}{\sqrt{\mathrm{Var} X}}$ ungefär $\mathrm{N}(0,1)$-fördelad.

Tvådimensionella slumpvariabler och fördelningar

- *Ifall Ω är ett utfallsrum på vilket en sannolikhetsfunktion är definierad så är $(X, Y) : \Omega \to \mathbb{R}^2$ en tvådimensionell slumpvariabel med fördelningsfunktion $F_{XY}(x, y) = \Pr(X \le x, Y \le y)$ förutsatt att $\{\,\omega \in \Omega : X(\omega) \le x, Y(\omega) \le y\,\}$ är en mängd för vilken sannolikheten är definierad.*
- *En tvådimensionell slumpvariabel (X, Y) är diskret och den har frekvensfunktionen $f_{XY}(x, y)$ ifall $f_{XY}(x, y) \ge 0$ för alla x och y, $\sum_x \sum_y f_{XY}(x, y) = 1$ och $\Pr(X = x, Y = y) = f_{XY}(x, y)$ för alla x och y.*
- *En tvådimensionell slumpvariabel (X, Y) är kontinuerlig och har täthetsfunktion $f_{XY}(x, y)$ ifall $f_{XY}(x, y) \ge 0$ för alla x och y, $\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{XY}(x, y)\,dx\,dy = 1$ och $F_{XY}(x, y) = \int_{-\infty}^{x}\int_{-\infty}^{y} f_{XY}(s, t)\,dt\,ds$ för alla x och y.*

Marginalfördelningar

- *Ifall $f_{XY}(x, y)$ är frekvensfunktionen för den diskreta tvådimensionella slumpvariabeln (X, Y) så är marginalfrekvensfunktionerna för slumpvariablerna X och Y*

$$f_X(x) = \Pr(X = x) = \sum_y f_{XY}(x, y) \quad \text{och}$$

$$f_Y(y) = \Pr(Y = y) = \sum_x f_{XY}(x, y).$$

- *Ifall $f_{XY}(x, y)$ är täthetsfunktionen för den kontinuerliga tvådimensionella slumpvariabeln (X, Y) så är marginaltäthetsfunktionerna för slumpvariablerna X och Y*

$$f_X(x) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,dy \quad \text{och} \quad f_Y(y) = \int_{-\infty}^{\infty} f_{XY}(x, y)\,dx.$$

Oberoende slumpvariabler

Om (X, Y) är en diskret tvådimensionell slumpvariabel eller en kontinuerlig tvådimensionell slumpvariabel med täthetsfunktion så gäller

$$X \text{ och } Y \text{ är oberoende} \quad \Leftrightarrow \quad f_{XY}(x, y) = f_X(x) f_Y(y).$$

Om X och Y är oberoende och f och g är mätbara funktioner så gäller

$$\mathrm{E}(f(X)g(Y)) = \mathrm{E}(f(X))\mathrm{E}(g(Y)).$$

Kovarians och korrelation

- *Kovarians:* $\mathrm{Cov}(X, Y) = \mathrm{E}\big((X - \mathrm{E}(X))(Y - \mathrm{E}(Y))\big) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$.
- *Korrelation:* $\mathrm{Cor}(X, Y) = \dfrac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}$, $-1 \le \mathrm{Cor}(X, Y) \le 1$.
- *Om X och Y är oberoende så är* $\mathrm{Cov}(X, Y) = \mathrm{Cor}(X, Y) = 0$.

Betingade fördelningar

- *Ifall (X, Y) är en diskret tvådimensionell slumpvariabel med frekvensfunktion $f_{XY}(x, y)$ eller en kontinuerlig tvådimensionell slumpvariabel med täthetsfunktion $f_{XY}(x, y)$ så är*

$$f_{Y|X}(y|x) = \frac{f_{XY}(x, y)}{f_X(x)}, \quad \text{då} \quad f_X(x) > 0,$$

  den betingande frekvensfunktionen respektive täthetsfunktionen för Y givet $X = x$.
- $\mathrm{E}(Y|X = x) = \begin{cases} \sum_y y f_{Y|X}(y|x) & \textit{eller} \\ \int_{-\infty}^{\infty} y f_{Y|X}(y|x)\,\mathrm{d}y \end{cases}$
- $\mathrm{E}(Y|X)$ *är en slumpvariabel så att* $\mathrm{E}(Y|X)(\omega) = \mathrm{E}(Y|X = X(\omega))$ *då* $\omega \in \Omega$.
- $\mathrm{E}(\mathrm{E}(Y|X)) = \mathrm{E}(Y)$.

Den tvådimensionella normalfördelningen

Slumpvariabeln (X, Y) sägs vara normalfördelad om varje linär kombination $\alpha X + \beta Y$ är normalfördelad och slumpvariabeln (X, Y) har då, ifall $\rho = \mathrm{Cor}(X, Y) \neq \pm 1$, $\sigma_X^2 = \mathrm{Var}(X) > 0$ och $\sigma_Y^2 = \mathrm{Var}(Y) > 0$, en täthetsfunktion med formen

$$f_{XY}(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}\left(\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - \frac{2\rho(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y}\right)}.$$

Då är $X \sim \mathrm{N}(\mu_X, \sigma_X^2)$, $Y \sim \mathrm{N}(\mu_Y, \sigma_Y^2)$ *och*

$$f_{Y|X}(y|x) = \frac{1}{\sqrt{2\pi}\sigma_Y\sqrt{1-\rho^2}} e^{-\frac{1}{2}\left(\frac{y-\mu_Y-\rho\frac{\sigma_Y}{\sigma_X}(x-\mu_X)}{\sigma_Y\sqrt{1-\rho^2}}\right)^2}$$

så att $(Y|X = x) \sim \mathrm{N}\left(\mu_Y + \rho\frac{\sigma_Y}{\sigma_X}(x - \mu_X), (1-\rho^2)\sigma_Y^2\right)$.
Om $\rho = 0$ är X och Y oberoende.

Obs!

Observera att om (X, Y) är en slumpvariabel så att X är normalfördelad och Y är normalfördelad så behöver inte (X, Y) nödvändigtvis vara normalfördelad men om tex. X och Y är oberoende så är också (X, Y) normalfördelad.

Ett annat uttryck för täthetsfunktionen då (X, Y) är normalfördelad

Ett annat sätt att skriva täthetsfunktionen som gör det lättare att generalisera till det m-dimensionella fallet är

$$f_{XY}(x, y) = \frac{1}{\sqrt{2\pi}^2\sqrt{\det(\mathbf{\Sigma})}} e^{-\frac{1}{2}(\mathbf{z}-\boldsymbol{\mu})^{\mathsf{T}}\mathbf{\Sigma}^{-1}(\mathbf{z}-\boldsymbol{\mu})}, \quad \mathbf{z} = \begin{bmatrix} x \\ y \end{bmatrix}, \quad \boldsymbol{\mu} = \begin{bmatrix} \mu_X \\ \mu_Y \end{bmatrix},$$

där $\mathbf{\Sigma} = \begin{bmatrix} \sigma_X^2 & \rho\sigma_X\sigma_Y \\ \rho\sigma_X\sigma_Y & \sigma_Y^2 \end{bmatrix}$, *så att* $\mathbf{\Sigma}^{-1} = \frac{1}{1-\rho^2}\begin{bmatrix} \frac{1}{\sigma_X^2} & -\frac{\rho}{\sigma_X\sigma_Y} \\ -\frac{\rho}{\sigma_X\sigma_Y} & \frac{1}{\sigma_X^2} \end{bmatrix}$ *och*
$\det(\mathbf{\Sigma}) = \sigma_X^2\sigma_Y^2(1-\rho^2)$.