Kapitel 2

Vidare Repetition

Marginalfördelning. För (ξ,η) en $\mathbb{R}^2$-värd s.v. med frekvensfunktion $f_{(\xi,\eta)}(x,y)$ har ξ frekvensfunktion

$$f_\xi(x) = \sum_y f_{(\xi,\eta)}(x,y)$$

om η diskret (och motsvarande integral om kontinuerlig). För varje värde x som ξ kan anta, så summerar vi alltså upp över alla värden som η kan anta då $\xi = x$.

Detta generaliseras även till $\mathbb{R}^n$-värda s.v. $\xi = (\xi_1,\ldots,\xi_n)$ med given frekvensfunktion och vi vill veta frekvensfunktionen för en av ξ_k eller för en sammansättning, t.ex (ξ_1,ξ_2).

Väntevärdet för $\mathbb{R}^n$-värd s.v. $\xi = (\xi_1,\ldots,\xi_n)$ definieras som en kolonnvektor

$$\mathbf{E}\{\xi\} = \begin{pmatrix}\mathbf{E}\{\xi_1\} & \mathbf{E}\{\xi_2\} & \ldots & \mathbf{E}\{\xi_n\}\end{pmatrix}^T$$

där väntevärdena $\mathbf{E}\{\xi_k\}$ beräknas som ett “vanligt” väntevärde av en $\mathbb{R}$-värd s.v. med hjälp av respektive marginalfrekvensfunktion f_{ξ_k}.

Notera att **väntevärdet för en linjärkombination av s.v.** (Sats 0.13) är

$$\mathbf{E}\{\sum_{k=1}^n a_k\,\xi_k\} = \sum_{k=1}^n a_k\,\mathbf{E}\{\xi_k\}$$

där $a_1,\ldots,a_n \in \mathbb{R}$ vilket följer av att

$$\mathbf{E}\{g(\xi)\} = \int_{x\in\mathbb{R}^n} g(x) f_\xi(x)\,dx$$

för ξ en $\mathbb{R}^n$-värd kontinuerlig s.v. och $g : \mathbb{R}^n \mapsto \mathbb{R}$, tillsammans med linjäriteten hos integraler.

Kovariansen mellan två $\mathbb{R}$-värda s.v. ξ och η definieras

$$\mathbf{Cov}\{\xi,\eta\} = \mathbf{E}\big\{(\xi - \mathbf{E}\{\xi\})(\eta - \mathbf{E}\{\eta\})\big\}$$

och **variansen** av ξ definieras

$$\mathbf{Var}\{\xi\} = \mathbf{Cov}\{\xi,\xi\} = \mathbf{E}\big\{(\xi - \mathbf{E}\{\xi\})^2\big\}$$

Notera att $\mathbf{Cov}\{\xi,\eta\} = \mathbf{Cov}\{\eta,\xi\}$, dvs kovariansen är **symmetrisk**.

Om $\mathbf{Cov}\{\xi,\eta\} = 0$ sägs ξ och η vara **okorrelerade**.

Kovariansen av två linjärkombinationer (Sats 0.15) $\sum_{k=1}^{n} a_k\,\xi_k$ och $\sum_{l=1}^{m} b_l\,\eta_l$ är

$$\mathbf{Cov}\{\sum_{k=1}^{n} a_k\,\xi_k, \sum_{l=1}^{m} b_l\,\eta_l\} = \sum_{k=1}^{n}\sum_{l=1}^{m} a_k b_l\,\mathbf{Cov}\{\xi_k,\eta_l\}$$

Kom-ihåg-regel: Jämför med multiplikation av två summor, t.ex

$$(\xi_1+\xi_2)(\eta_1+\eta_2) = \xi_1\eta_1+\xi_1\eta_2+\xi_2\eta_1+\xi_2\eta_2$$

med motsvarande för kovarianser; vi får enligt Sats 0.15:

$$\mathbf{Cov}\{\xi_1+\xi_2,\eta_1+\eta_2\} = \mathbf{Cov}\{\xi_1,\eta_1\}+\mathbf{Cov}\{\xi_1,\eta_2\}+\mathbf{Cov}\{\xi_2,\eta_1\}+\mathbf{Cov}\{\xi_2,\eta_2\}$$

Om $\eta = \xi$ i Sats 0.15 får vi Sats 0.16:

$$\mathbf{Cov}\{a^T\xi, b^T\xi\} = \mathbf{Cov}\{\sum_{k=1}^{n} a_k\,\xi_k, \sum_{l=1}^{n} b_l\,\xi_l\} = \sum_{k=1}^{n}\sum_{l=1}^{n} a_k b_l \mathbf{Cov}\{\xi_k,\xi_l\} = a^T\mathbf{Var}\{\xi\}b$$

där $\mathbf{Var}\{\xi\}$ kallas **variansmatrisen**, som är en $n\times n$ matris med element $\mathbf{Cov}\{\xi_k,\xi_l\}$.

Ytterligare generaliserat ger detta också, för $\eta = A\xi$ och $A\in\mathbb{R}^{m\times n}$,

$$\mathbf{Var}\{\eta\} = \mathbf{Var}\{A\xi\} = A\mathbf{Var}\{\xi\}A^T$$

Faktum: Varje variansmatris V för någon $\mathbb{R}^n$-värd s.v. måste vara <u>symmetrisk</u> och <u>positivt semidefinit</u>. Symmetrin följer av att $V_{i,j} = \mathbf{Cov}\{\xi_i,\xi_j\} = \mathbf{Cov}\{\xi_j,\xi_i\} = V_{j,i}$ och positivt semidefinit av observationen att för alla $a\in\mathbb{R}^n$

$$a^T\mathbf{Var}\{\xi\}\,a = [\text{Sats } 0.16] = \mathbf{Cov}\{a^T\xi, a^T\xi\} = \mathbf{Var}\{a^T\xi\} \geq 0$$

eftersom en varians aldrig är negativ (ty väntevärde av en kvadrat). Att $a^T\mathbf{Var}\{\xi\}a\geq 0$ för alla $a\in\mathbb{R}^n$ är just definitionen på en positivt semidefinit matris.

Ytterligare en viktig slutsats från ovanstående räkningar med kovarians mellan linjärkombinationer är att om $\xi_1,\ldots,\xi_n$ är okorrelerade så är

$$\mathbf{Var}\{\sum_{k=1}^{n} a_k\xi_k\} = \sum_{k=1}^{n} a_k^2\,\mathbf{Var}\{\xi_k\}.$$

Oberoende stokastiska variabler

Kom ihåg: Händelserna $A,B\subseteq\Omega$ är oberoende om $\mathbf{P}(A\cap B) = \mathbf{P}(A)\mathbf{P}(B)$.

För stokastiska variabler gäller motsvarande: ξ och η ($\mathbb{R}^n$- resp. $\mathbb{R}^m$-värda s.v.) är oberoende om

$$\mathbf{P}(\xi\in A,\eta\in B) = \mathbf{P}(\xi\in A)\mathbf{P}(\eta\in B)$$

för alla $A \subseteq \mathbb{R}^n$ och $B \subseteq \mathbb{R}^m$ (där kommat i sannolikheten till vänster tolkas som $\cap$).

Vi måste alltså titta på alla möjliga händelser i ξ:s och η:s värdemängder. Detta är generellt svårt, helt enkelt därför att antalet möjliga mängder är oändligt många. Som tur är finns det andra ekvivalenta villkor som är enklare att verifiera:

De stokastiska variablerna $\xi : \Omega \mapsto \mathbb{R}^n$ och $\eta : \Omega \mapsto \mathbb{R}^m$ är oberoende om och endast om

$$f_{(\xi,\eta)}(x, y) = f_\xi(x) f_\eta(y)$$

för alla $x \in \mathbb{R}^n$ och alla $y \in \mathbb{R}^m$.

En bra grundregel för oberoende är att

oberoende faktoriserar

Här kommer ytterligare ett exempel på det:

För oberoende $\mathbb{R}$-värda stokastiska variabler ξ och η har vi

$$\mathbf{E}\{\xi\eta\} = \mathbf{E}\{\xi\}\mathbf{E}\{\eta\}$$

Vi har således också speciellt att $\mathbf{Cov}\{\xi,\eta\} = \mathbf{E}\{\xi\eta\} - \mathbf{E}\{\xi\}\mathbf{E}\{\eta\} = 0$ för ξ och η oberoende. Att kovariansen är noll kallas att variablerna är okorrelerade.

Notera att detta förhållande mellan väntevärdena är mycket svagt och det omvända gäller inte allmänt, dvs okorrelerat innebär inte, generellt sett, oberoende.

Ett exempel som är värt att lägga på minnet är följande övning från kursboken:

Övning 2.5: *Är ξ och η oberoende och/eller okorrelerade då (ξ,η) är en $\mathbb{R}^2$-värd s.v. med värdemängd $\Omega_{(\xi,\eta)} = \{(0,0),(\pm1,\pm1)\}$ och $f_{(\xi,\eta)}(0,0) = 1/2$ och $f_{(\xi,\eta)}(\pm1,\pm1) = 1/8$?*

Lösning: *De är oberoende om $f_{(\xi,\eta)}(x,y) = f_\xi(x)f_\eta(y)$ för alla $(x,y) \in \Omega_{(\xi,\eta)}$ enligt satsen ovan. För att kontrollera detta måste vi först beräkna marginalerna f_ξ och f_η. Enkla räkningar ger*

$$f_\xi(x) = \sum_y f_{(\xi,\eta)}(x,y) = \begin{cases} 1/2 & \textit{för } x = 0 \\ 1/4 & \textit{för } x = \pm1 \end{cases}$$

och samma för f_η.

Vi har nu att $f_{(\xi,\eta)}(0,0) = 1/2 \neq 1/4 = f_\xi(0)f_\eta(0)$. så ξ och η är inte oberoende (det räcker ju att hitta ett exempel på att faktoriseringen inte håller).

Om man tänker efter så bör detta vara uppenbart eftersom om vi vet att $\eta = 0$ så måste ju $\xi = 0$; vi får alltså information om vad ξ kan vara genom att tala om vad η är, vilket inte skall gälla om de var oberoende.

Okorrelerade då? Vi behöver kontrollera om $\mathbf{E}\{\xi\eta\} = \mathbf{E}\{\xi\}\mathbf{E}\{\eta\}$ Vi har

$$\begin{aligned}\mathbf{E}\{\xi\eta\} = \sum_{(x,y)\in\Omega_{(\xi,\eta)}} xyf_{(\xi,\eta)}(x,y) = 0 \cdot f_{(\xi,\eta)}(0,0) + \\ + 1 \cdot \big(f_{(\xi,\eta)}(1,1) + f_{(\xi,\eta)}(-1,-1)\big) + (-1) \cdot \big(f_{(\xi,\eta)}(1,-1) + f_{(\xi,\eta)}(-1,1)\big) = 0\end{aligned}$$

och vi ser enkelt att $\mathbf{E}\{\xi\} = \mathbf{E}\{\eta\} = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{4} + (-1) \cdot \frac{1}{4} = 0$ och drar sålunda slutsatsen att de är okorrelerade eftersom $\mathbf{E}\{\xi\eta\} = 0 = \mathbf{E}\{\xi\}\mathbf{E}\{\eta\}$.

Detta är alltså ett exempel på två stokastiska variabler som är okorrelerade men inte oberoende och är därmed ett exempel på att okorrelerat inte innebär oberoende.

Betingade stokastiska variabler

Kom ihåg att vi, för händelser $A, B \subseteq \Omega$ med $\mathbf{P}(B) > 0$, definierade den betingade sannolikheten för A givet att B inträffat som

$$\mathbf{P}(A|B) = \frac{\mathbf{P}(A \cap B)}{\mathbf{P}(B)}$$

Något liknande måste ju rimligtvis även finnas för stokastiska variabler.

Med $f_{(\xi|\eta)}(x|y)$ menar vi frekvensfunktionen för ξ givet att $\eta = y$ och definierar den som

$$f_{(\xi|\eta)}(x|y) = \begin{cases} \frac{f_{(\xi,\eta)}(x,y)}{f_\eta(y)} & \text{för } f_\eta(y) > 0 \\ 0 & \text{för } f_\eta(y) = 0 \end{cases}$$

Det betingade väntevärdet för ξ givet $\eta = y$, definierar vi som

$$\mathbf{E}\{\xi\,|\eta = y\} = \begin{cases} \int x f_{(\xi|\eta)}(x|y)\,dx & \text{för } \xi \text{ kontinuerlig} \\ \sum x f_{(\xi|\eta)}(x|y) & \text{för } \xi \text{ diskret} \end{cases}$$

En användbar metod för att beräkna väntevärdet $\mathbf{E}\{\xi\}$ (om $\mathbf{E}\{\xi|\eta = y\}$ är enkel att beräkna), är

$$\mathbf{E}\{\xi\} = \begin{cases} \int \mathbf{E}\{\xi\,|\eta = y\}\, f_\eta(y)\,dy & \text{för } \eta \text{ kontinuerlig} \\ \sum \mathbf{E}\{\xi\,|\eta = y\}\, f_\eta(y) & \text{för } \eta \text{ diskret} \end{cases}$$

Med anknytning till Exempel 4 från Fö1 har vi följande.

Exempel: *Antag att vi får ett Po(λ)-fördelat antal lotter där vardera lott har en vinstchans på $p \in (0,1)$, oberoende av de andra lotterna.*

Låt η vara antalet vinstlotter och ξ det totala antalet lotter. Vi har att $\xi \sim Po(\lambda)$, dvs att $f_\xi(x) = \frac{\lambda^x}{x!}e^{-\lambda}$ enligt definitionen på Poissonfördelningen.

Givet ett visst utfall på ξ, låt oss säga $\xi = n$, har vi alltså n stycken lotter, vardera med vinstchans p, oberoende av de andra. Men detta innebär ju att fördelningen för η givet att $\xi = n$ är Bin(n, p). För binomialfördelade s.v. med parametrar n och p har vi att väntevärdet är np; alltså är

$$\mathbf{E}\{\eta\,|\xi = n\} = np$$

Om vi använder satsen ovan för att beräkna $\mathbf{E}\{\eta\}$ så får vi

$$\mathbf{E}\{\eta\} = \sum_{n=0}^{\infty} \mathbf{E}\{\eta\,|\xi = n\} f_\xi(n) = \sum_{n=0}^{\infty} np\, f_\xi(n) = p\sum_{n=0}^{\infty} n f_\xi(n) = p\mathbf{E}\{\xi\} = p\lambda$$

eftersom den sista summan är inget annat än definitionen för väntevärdet av en diskret s.v. ξ, samt att väntevärdet för en Po(λ)-fördelad s.v. är λ.