Föreläsning 5

Thomas Önskog

8/11 2017

Flerdimensionella stokastiska variabler

Vi avslutade förra föreläsningen med att definiera diskreta tvådimensionella s.v. och definierar nu även kontinuerliga tvådimensionella s.v.

Definition. *En tvådimensionell s.v. (X, Y) sägs vara* ***kontinuerlig*** *om det existerar en icke-negativ funktion $f_{X,Y}$, kallad den* ***simultana täthetsfunktionen*** *för (X, Y), sådan att*

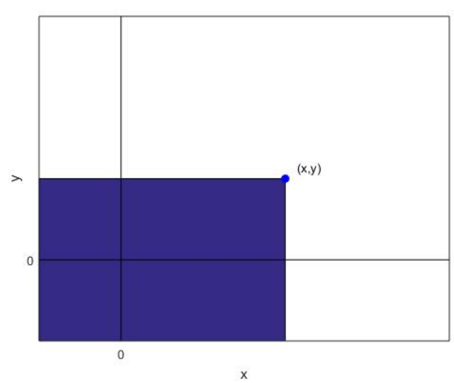
$$\underline{\mathbb{P}((X,Y) \in A) = \iint_A f_{X,Y}(x,y)dxdy}, \quad \text{för alla mängder } A \subset \mathbb{R}^2.$$

Notera att $\iint_{\mathbb{R}^2} f_{X,Y}(x,y)dxdy = 1$.

Både diskreta och kontinuerliga tvådimensionella s.v. kan beskrivas med en fördelningsfunktion.

Definition. *Den* ***simultana fördelningsfunktionen*** *för en tvådimensionell s.v. (X, Y) ges av*

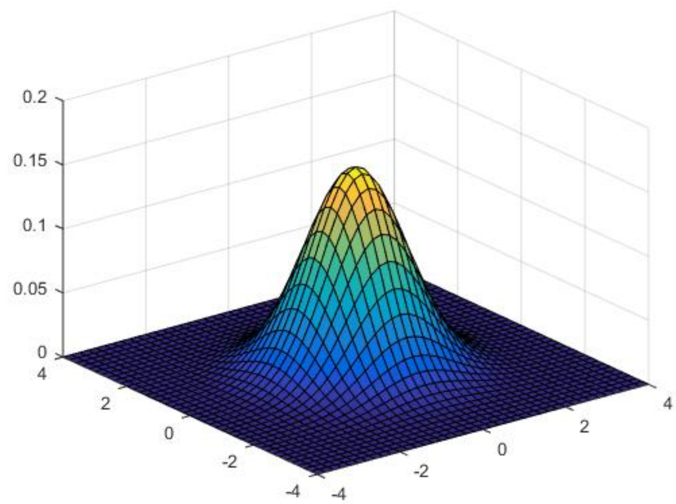
$$\underline{F_{X,Y}(x,y) = \mathbb{P}((X \le x) \cap (Y \le y))}.$$



Värdet på den simultana fördelningsfunktionen i punkten (x, y) är lika med sannolikheten att värdet på den s.v. (X, Y) ligger i det blåa området i figuren. Från definitionerna får vi följande samband.

$$\underline{F_{X,Y}(x,y) = \sum_{j \le x} \sum_{k \le y} p_{X,Y}(j,k)} \qquad \text{och} \qquad \underline{F_{X,Y}(x,y) = \int_{-\infty}^{x} \Big(\int_{-\infty}^{y} f_{X,Y}(u,v)dv \Big) du}.$$

för diskreta respektive kontinuerliga tvådimensionella s.v. Bilden nedan visar den simultana täthetsfunktionen för en tvådimensionall normalfördelning. Båda komponenterna i en tvådimensionell s.v. (X, Y) är endimensionella s.v. Komponenternas fördelningar kallas marginalfördelningar.



Definition. *För diskreta tvådimensionella s.v.* (X, Y) *så ges de* ***marginella sannolikhetsfunktionerna*** *för* X *och* Y *av*

$$p_X(j) = \sum_k p_{X,Y}(j,k) \qquad \textit{respektive} \qquad p_Y(k) = \sum_j p_{X,Y}(j,k).$$

För kontinuerliga tvådimensionella s.v. (X, Y) *så ges de* ***marginella täthetsfunktionerna*** *för* X *och* Y *av*

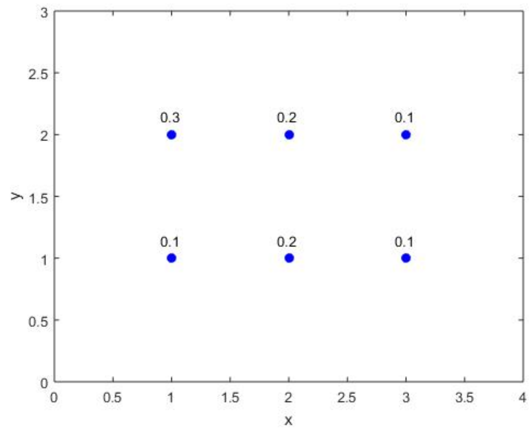
$$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy \qquad \textit{respektive} \qquad f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dx.$$

Exempel. *Figuren överst på nästa sida visar värdena för den simultana sannolikhetsfunktionen för en tvådimensionell s.v.* (X, Y)*. För denna s.v. har exempelvis den marginella sannolikhetsfunktionen för* X *i punkten* $x = 2$ *värdet*

$$p_X(2) = \sum_k p_{X,Y}(2,k) = 0.2 + 0.2 = 0.4,$$

och den marginella sannolikhetsfunktionen för Y *i punkten* $y = 1$ *har värdet*

$$p_Y(1) = \sum_j p_{X,Y}(j,1) = 0.1 + 0.2 + 0.1 = 0.4.$$



Oberoende stokastiska variabler

Vi har tidigare definierat två händelser A och B som oberoende om villkoret $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$ är uppfyllt. På liknande sätt sägs två s.v. X och Y vara **oberoende** om

$$\underline{\mathbb{P}((X \in C) \cap (Y \in D)) = \mathbb{P}(X \in C)\mathbb{P}(Y \in D)}, \quad \text{för alla mängder } C, D \in \mathbb{R}.$$

Som ett exempel på oberoende s.v. ska vi undersöka fördelningen för största och minsta värdet av två oberoende s.v. X och Y. Låt först $Z = \max\{X, Y\}$. Notera att $Z \le u$ om och endast om både $X \le u$ och $Y \le u$ gäller. Detta ger

$$F_Z(u) = \mathbb{P}(Z \le u) = \mathbb{P}((X \le u) \cap (Y \le u)) = \mathbb{P}(X \le u)\mathbb{P}(Y \le u) = F_X(u)F_Y(u),$$

där vi i näst sista steget använde att X och Y är oberoende. Låt nu $W = \min\{X, Y\}$. Då gäller $W > v$ om och endast om både $X > v$ och $Y > v$ gäller. Detta ger

$$\begin{aligned} F_W(v) &= 1 - \mathbb{P}(W > v) = 1 - \mathbb{P}((X > v) \cap (Y > v)) = 1 - \mathbb{P}(X > v)\mathbb{P}(Y > v) \\ &= 1 - (1 - F_X(v))(1 - F_Y(v)), \end{aligned}$$

där vi återigen använde att X och Y är oberoende i näst sista steget. Formlerna ovan kan lätt generaliseras till maximum och minimum av n s.v.

Exempel. *En motor har fyra cylindrar, vars livslängder X_i är oberoende och $Exp(1/7)$-fördelade. Låt T beteckna den tidpunkt då motorns kapacitet reduceras till följd av att en av cylindrarna går sönder, dvs $T = min(X_1, X_2, X_3, X_4)$. Från formlerna som vi nyss härledde följer att*

$$F_T(t) = 1 - \prod_{i=1}^{4}(1 - F_{X_i}(t)) = 1 - \prod_{i=1}^{4}(1 - (1 - e^{-t/7})) = 1 - (e^{-t/7})^4 = 1 - e^{-4t/7},$$

så $T \in Exp(4/7)$. Generellt gäller det att minimum av n oberoende $Exp(\lambda)$-fördelade s.v. är $Exp(n\lambda)$-fördelat.

Följande sats är användbar för att avgöra om två s.v. är oberoende.

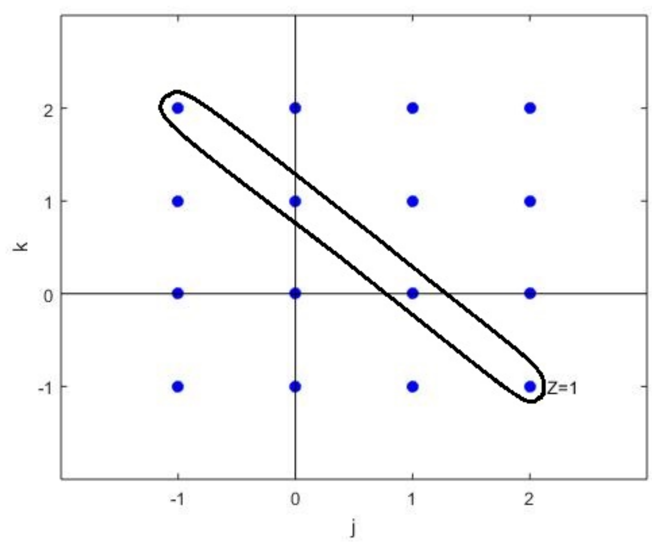
Sats. *Utsagorna 1-3 är ekvivalenta.*

1. *De s.v.* X *och* Y *är oberoende.*
2. $F_{X,Y}(x,y) = F_X(x)F_Y(y)$, *för alla* $x, y \in \mathbb{R}$.
3. $p_{X,Y}(j,k) = p_X(j)p_Y(k)$, *för alla* j, k *(om* X *och* Y *är diskreta s.v.)*
   $f_{X,Y}(x,y) = f_X(x)f_Y(y)$, *för alla* $x, y \in \mathbb{R}$ *(om* X *och* Y *är kontinuerliga s.v.)*

Största och minsta värdet av två oberoende s.v. X och Y kan ses som funktioner av en tvådimensionell s.v. (X, Y). En annan sådan funktion är $Z = X + Y$. Om X och Y är diskreta, oberoende s.v., så blir sannolikhetsfunktionen för $Z = X + Y$

$$p_Z(z) = \mathbb{P}(X + Y = z) = \sum_{j+k=z} p_{X,Y}(j,k) = \sum_j p_{X,Y}(j, z-j) = \sum_j p_X(j)p_Y(z-j),$$

där vi använde satsen ovan och oberoendet hos X och Y i sista steget. Formeln som vi härlett, $\underline{p_Z(z) = \sum_j p_X(j)p_Y(z-j)}$, kallas för faltningsformeln.



För kontinuerliga s.v. finns en direkt motsvarighet i formeln $\underline{f_Z(z) = \int_{-\infty}^{\infty} f_X(x)f_Y(z-x)dx}$ (se boken för härledningen).

Lägesmått och spridningsmått

När vi studerade deskriptiv statistik på föreläsning 1, så införde vi lägesmått och spridningsmått för datamängder. Även för sannolikhetsfördelningar är dylika mått av intresse. Det vanligaste lägesmåttet är väntevärdet och de vanligaste spridningsmåtten är variansen och standardavvikelsen.

Definition. *För en diskret s.v. X definieras **väntevärdet** som*

$$\underline{E(X) = \sum_j j p_X(j)},$$

och för en kontinuerlig s.v. X så definieras det som

$$\underline{E(X) = \int_{-\infty}^{\infty} x f_X(x) dx}.$$

Exempel. *Väntevärdet för en $Po(\mu)$-fördelad s.v. X ges av*

$$E(X) = \sum_{j=0}^{\infty} j \frac{\mu^j}{j!} e^{-\mu} = \sum_{j=1}^{\infty} \frac{\mu^j}{(j-1)!} e^{-\mu} = \mu e^{-\mu} \sum_{j=1}^{\infty} \frac{\mu^{j-1}}{(j-1)!} = \mu e^{-\mu} \underbrace{\sum_{k=0}^{\infty} \frac{\mu^k}{k!}}_{=e^{\mu}} = \mu$$

Exempel. *Väntevärdet för en $Exp(\lambda)$-fördelad s.v. X ges av*

$$\begin{aligned} E(X) &= \int_0^{\infty} x \lambda e^{-\lambda x} dx = [\textit{Partiell integration}] = \underbrace{[x(-e^{-\lambda x})]_0^{\infty}}_{=0} - \int_0^{\infty} -e^{-\lambda x} dx \\ &= [-\frac{1}{\lambda} e^{-\lambda x}]_0^{\infty} = \frac{1}{\lambda} \end{aligned}$$

Väntevärdet för en funktion g av en s.v. X ges av $\underline{E(g(X)) = \sum_j g(j) p_X(j)}$ om X är diskret och av $\underline{E(g(X)) = \int_{-\infty}^{\infty} g(x) f_X(x) dx}$ om X är kontinuerlig. Vi använder dessa definitioner i följande definition.

Definition. *Låt μ beteckna väntevärdet av en s.v. X. **Variansen** av X definieras som*

$$\underline{V(X) = E((X-\mu)^2)}.$$

*Om X är diskret, så gäller $\underline{V(X) = \sum_j (j-\mu)^2 p_X(j)}$ och om X är kontinuerlig, så gäller $\underline{V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f_X(x) dx}$. I båda fallen definieras **standardavvikelsen** av X som*

$$\underline{D(X) = \sqrt{V(X)}}.$$

Vid beräkning av varianser och standardavvikelser har vi ofta användning av följande sats.

Sats. *För alla s.v. X, så gäller $\underline{V(X) = E(X^2) - (E(X))^2}$.*

Bevis. Vi bevisar satsen för diskreta s.v. X, men beviset är analogt för kontinuerliga s.v. Låt μ beteckna $E(X)$ i beviset. Vi utvecklar $V(X)$

$$\begin{aligned} V(X) &= \sum_j (j-\mu)^2 p_X(j) = \sum_j (j^2 - 2\mu j + \mu^2) p_X(j) \\ &= \sum_j j^2 p_X(j) - 2\mu \underbrace{\sum_j j p_X(j)}_{=E(X)=\mu} + \mu^2 \underbrace{\sum_j p_X(j)}_{=1} = E(X^2) \underbrace{-2\mu \cdot \mu + \mu^2}_{=-\mu^2=-(E(X))^2}. \end{aligned}$$

$\square$

Vi använder satsen för att bestämma variansen av en Exp(λ)-fördelad s.v. X.

Exempel. *För $X \in Exp(\lambda)$, så gäller*

$$\begin{aligned} E(X^2) &= \int_0^\infty x^2 \lambda e^{-\lambda x} dx = [\textit{Partiell integration}] = \underbrace{[x^2(-e^{-\lambda x})]_0^\infty}_{=0} - \int_0^\infty 2x(-e^{-\lambda x}) dx \\ &= \frac{2}{\lambda} E(X) = \frac{2}{\lambda^2}, \end{aligned}$$

vilket ger

$$V(X) = E(X^2) - (E(X))^2 = \frac{2}{\lambda^2} - \left(\frac{1}{\lambda}\right)^2 = \frac{1}{\lambda^2},$$

och därmed $D(X) = 1/\lambda$.