Föreläsning 5: Flerdimensionella s.v.

Johan Thim (johan.thim@liu.se)

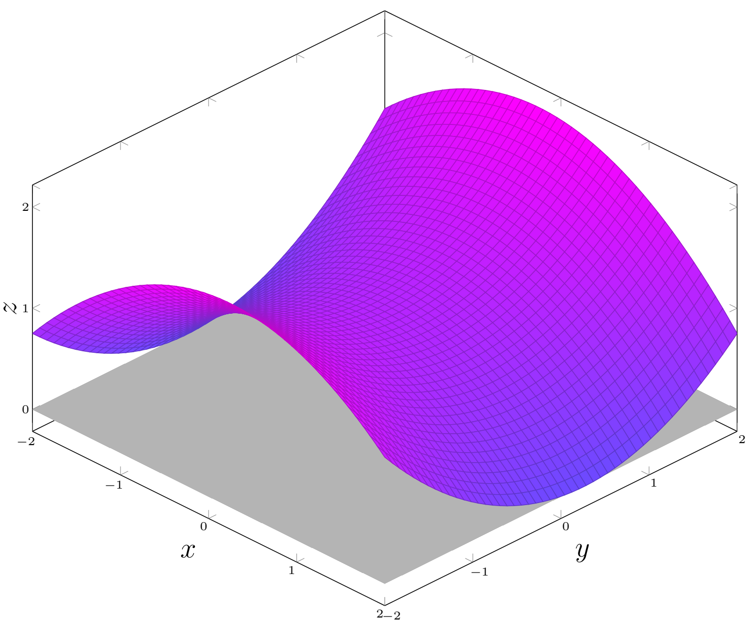
January 15, 2022

1 Flervariabelanalys

Precis som i linjär algebra så betecknar vi vektorer enligt $\boldsymbol{x} = (x_1, x_2, \ldots, x_n) \in \mathbf{R}^n$. Har vi bara två eller tre variabler använder vi ofta uttrycken (x, y) respektive (x, y, z) istället av tradition.

1.1 Funktioner av flera variabler

Så vi behöver en gnutta flervariabelanalys för att kunna hantera det som komma skall. För att komma åt de bitar vi behöver så måste vi generalisera integralbegreppet till högre dimensioner. Som exempel betraktar vi situationen med en funktion som beror på två variabler (säg x och y) och som ger ett reellt tal som "utdata." Mer kompakt uttryckt så skriver vi att $f : \mathbf{R}^2 \to \mathbf{R}$. Typiskt för oss är att vi tänker oss att $z = f(x, y)$, där vi tolkar funktionsvärdet som höjden ovanför (eller under om vi har negativt tecken) xy-planet. Skulle man försöka rita en figur skulle vi kunna få något enligt nedan.



Det gråa området där $z = 0$ (under själva grafen) är det område där (x, y) varierar över. Generellt sett så behöver inte detta vara en rektangel utan kan vara i princip vilket hyfsat snällt två-dimensionellt objekt som helst. I varje punkt (x, y) i detta område har vi alltså en "höjd" $z = f(x, y)$. Här kan vi givetvis analysera hur denna funktion beter sig och introducera begrepp som kontinuitet, derivata etc i flera variabler, men det blir lite för omfattande. Vad vi däremot kommer behöva är ett volymsbegrepp. Tanken är att vi skulle vilja räkna ut volymen mellan xy-planet och ytan $z = f(x, y)$.

1.2 Integration av funktioner av flera variabler

Om vi jämför med det endimensionella fallet, där $y = f(x)$, så approximerade vi arean av ett område $0 \leq y \leq f(x)$ med rektanglar och summerade arean av dessa för att konstruera Riemannintegralen av $f(x)$ (bläddra tillbaka till kapitel 6 i envariabelboken). Man kan göra motsvarande manöver i högre dimensioner, men istället för rektanglar får vi nu summera volymen på rätblock (staplar) som får mindre och mindre tvärsnittsarea. För en funktion av två variabler, där $z = f(x, y)$, så ter sig detta ganska naturligt rent geometriskt. Vi staplar in legobitar under ytan så högt som möjligt utan att sticka ut för mycket från ytan $z = f(x, y)$. Vi betecknar denna volym med

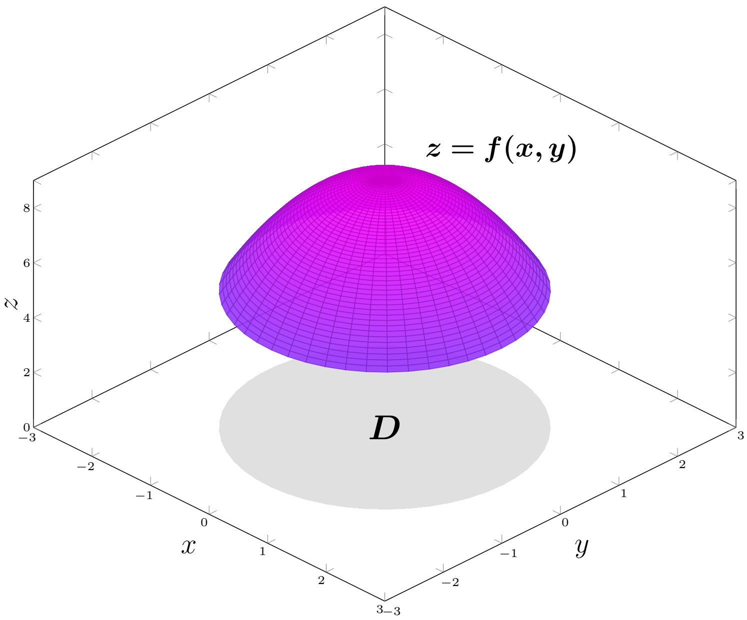
$$V = \iint_D f(x, y)\, dxdy$$

där $D \subset \mathbf{R}^2$ är det område som vi låter (x, y) variera över.

Detta generaliserar sedan naturligt till $\mathbf{R}^n$ (med skillnaden att det nu blir aningen svårare att visualisera). Konstruktionen kräver egentligen en hel del bevis (samt egentligen ett annat integralbegrepp då Riemannintegralen är lite problematisk), men vi skippar det och fokuserar istället på hur vi rent analytiskt hanterar integralerna om vi vill räkna ut ”volymen.”

1.3 Iteration i 2D

Låt oss fokusera på fallet i två dimensioner.

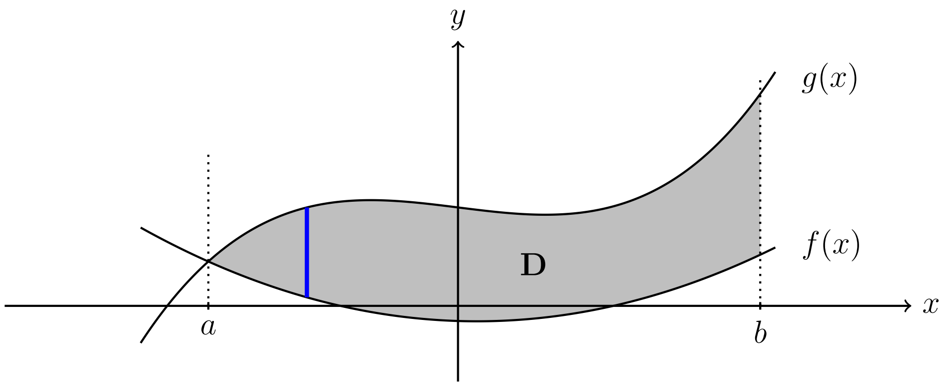


Den funktion som integreras ligger i en dimension ovanför (tänkt att du har en z-axel som pekar rakt upp ur sidan och att höjden ovanför sidan är funktionsvärdet $z = f(x, y)$ i puntken (x, y)). Typiskt skulle det kunna se ut enligt ovan. Volymen mellan ytan $z = f(x, y)$ och xy-planet när $(x, y) \in D$ betecknar vi nu med

$$V = \iint_D f(x, y)\, dxdy.$$

Hur räknar vi ut detta? Generellt så kan det vara svårt att hitta ett analytiskt uttryck, men om D uppfyller vissa krav kan vi *iterera* integralerna och räkna ut en i taget. Betrakta följande specialfall där vi kan stänga in y-koordinaten mellan funktioner av x:

$$D = \{(x, y) \in \mathbf{R}^2 : a \leq x \leq b,\ f(x) \leq y \leq g(x)\}.$$



Vi tänkar oss nu att vi styckar upp D i ”stolpar” som är parallella med y-axeln, så att om $a \leq x \leq b$ så täcker dessa stolpar precis området D (se den blåa linjen i figuren ovan). Det går i detta fall att visa att

$$\iint_D f(x, y)\, dxdy = \int_a^b \left(\int_{f(x)}^{g(x)} f(x, y)\, dy \right) dx$$

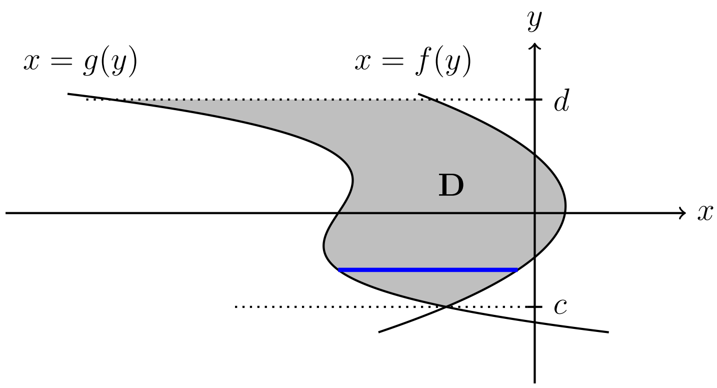
Motsvarande kan göras om området har formen

$$D = \{(x, y) \in \mathbf{R}^2 : c \leq y \leq d,\ f(y) \leq x \leq g(y)\}$$

så

$$\iint_D f(x, y)\, dxdy = \int_c^d \left(\int_{f(y)}^{g(y)} f(x, y)\, dx \right) dy.$$

Stolparna är nu parallella med x-axeln istället.



Notera speciellt att det är konstanter i gränserna på de yttre integralerna i båda fallen. Vi får *inte* ha några x eller y löst hängande när vi är färdiga. Skulle det inte gå att få till ett av fallen kan man behöva stycka upp området. Det fungerar utan problem: om $D = D_1 \cup D_2$ så gäller att

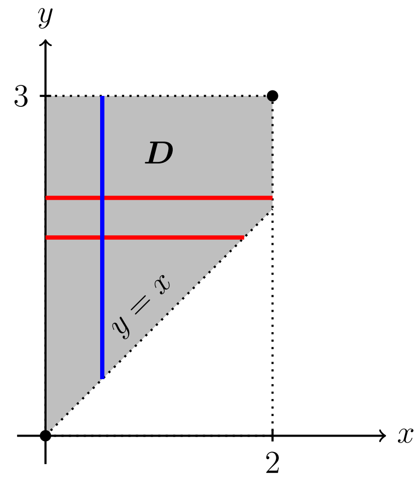
$$\iint_D f(x, y)\, dxdy = \iint_{D_1} f(x, y)\, dxdy + \iint_{D_2} f(x, y)\, dxdy,$$

så vi kan hantera en del av området i taget.

Exempel

Beräkna $\iint_D (xy+2)\,dxdy$ om D är den del av rektangeln med hörn i $(0,0)$ och $(2,3)$ där $y > x$.

Lösning. Låt oss börja med att rita en figur så vi ser vad D är.



Det skuggade området är det område D vi vill integrera över. Vi har två alternativ. Endera så itererar vi med de röda linjerna eller längs de blå. Vi ställer upp båda fallen som övning (de kommer alltid att ge samma svar).

Alternativ 1. Om vi väljer de röda linjerna ser vi att längden på dessa beter sig olika beroende på om $0 \le y \le 2$ eller om $2 < y \le 3$. Om $0 \le y \le 2$ så sträcker sig x från 0 till linjen $y = x$, så $0 \le x \le y$. Om $2 < y \le 3$ sträcker sig x från 0 till 2. Därför delar vi upp i två delar:

$$\begin{aligned}
\iint_D (xy+2)\,dxdy &= \int_0^2 \left(\int_0^y (xy+2)\,dx\right)\,dy + \int_2^3 \left(\int_0^2 (xy+2)\,dx\right)\,dy \\
&= \int_0^2 \left[\frac{x^2}{2}y + 2x\right]_{x=0}^{y} dy + \int_2^3 \left[\frac{x^2}{2}y + 2x\right]_{x=0}^{2} dy \\
&= \int_0^2 \left(\frac{y^3}{2} + 2y - 0\right) dy + \int_2^3 \left(\frac{2^2}{2}y + 4 - 0\right) dy \\
&= \left[\frac{y^4}{8} + y^2\right]_0^2 + \left[y^2 + 4y\right]_2^3 = \frac{2^4}{8} + 2^2 + 3^2 + 12 - (2^2+8) = 15.
\end{aligned}$$

Alternativ 2. Om vi väljer de blåa linjerna ser vi att för $0 \le x \le 2$ så sträcker sig y från linjen $y = x$ till $y = 3$. Vi behöver alltså inte dela upp integralen:

$$\begin{aligned}
\iint_D (xy+2)\,dxdy &= \int_0^2 \left(\int_x^3 (xy+2)\,dy\right)\,dx = \int_0^2 \left[x\frac{y^2}{2} + 2y\right]_{y=x}^{3} dx \\
&= \int_0^2 \left(\frac{9x}{2} + 6 - \left(\frac{x^3}{2} + 2x\right)\right) dx = \int_0^2 \left(\frac{5x - x^3}{2} + 6\right) dx \\
&= \left[\frac{5x^2}{4} - \frac{x^4}{8} + 6x\right]_0^2 = \frac{20}{4} - 2 + 12 = 15.
\end{aligned}$$

Vilket alternativ var enklast?

2 Flerdimensionella stokastiska variabler

Vi fokuserar på två-dimensionella variabler. Det är steget från en dimension till två som är det svåraste. Generaliseringar till högre dimensioner följer utan problem i de flesta fall.

Stokastisk variabel

Definition. En tvådimensionell stokastisk variabel är en reell-vektorvärd funktion (X, Y) (två komponenter) definierad på ett utfallsrum Ω. Alltså avbildar (X, Y) olika utfall på reella vektorer; $(X, Y)\colon \Omega \to \mathbf{R}^2$. Om (X, Y) bara antar ändligt eller uppräkneligt många värden så kallar vi (X, Y) för en diskret stokastisk variabel. Om varken X eller Y är diskret kallar vi (X, Y) för kontinuerlig.

Definitionen är analog med envariabelfallet. Observera dock följande: en situation som kan uppstå är att vi får ”halvdiskreta” variabler med ena variabeln diskret och den andra kontinuerlig! Inträffar inte ofta i denna kurs, men det kan vara värt att beakta.

Exempel

(i) Låt (X, Y) vara vikten X och längden Y hos en person. Variabeln är kontinuerlig och $\Omega = (X(\Omega), Y(\Omega)) = [0, \infty)^2 = [0, \infty) \times [0, \infty)$.

(ii) Låt (X, Y) vara resultaten av ett tärningskast (X) respektive ett myntkast (Y), där vi representerar krona med 1 och klave med -1. Variabeln är diskret och:

$$\begin{aligned} \Omega &= \{\, \text{⚀}, \text{⚁}, \text{⚂}, \text{⚃}, \text{⚄}, \text{⚅} \,\} \times \{\, \text{Krona}, \text{Klave} \,\}, \\ (X(\Omega), Y(\Omega)) &= \{\, 1, 2, 3, 4, 5, 6 \,\} \times \{\, -1, 1 \,\}. \end{aligned}$$

I det 2-dimensionella fallet är vi nu intresserade av delmängder i planet. För att kunna prata om en fördelningsfunktion introducerar vi mängden

$$C(x, y) = \{(u, v) \in \mathbf{R}^2 : u \leq x \text{ och } v \leq y\}.$$

Vi ser att $P((X, Y) \in C(x, y)) = P(X \leq x, Y \leq y)$ och gör följande definition.

Fördelningsfunktion

Definition. Funktionen $F_{X,Y}(x, y) = P(X \leq x, Y \leq y)$ kallas fördelningsfunktionen för den stokastiska variabeln (X, Y).

Om (X, Y) är diskret låter vi

$$p_{X,Y}(j, k) = P(X = j, Y = k).$$

Detta är **sannolikhetsfunktionen** för (X, Y). Fördelningsfunktionen ges då av

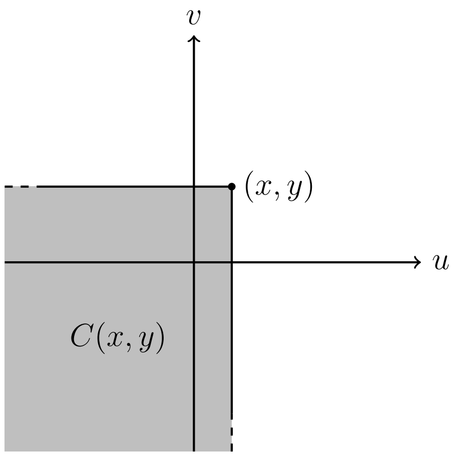
$$F_{X,Y}(x, y) = \sum_{j \leq x} \sum_{k \leq y} p_{X,Y}(j, k).$$

Om det finns en icke-negativ integrerbar funktion $f_{X,Y}$ så att

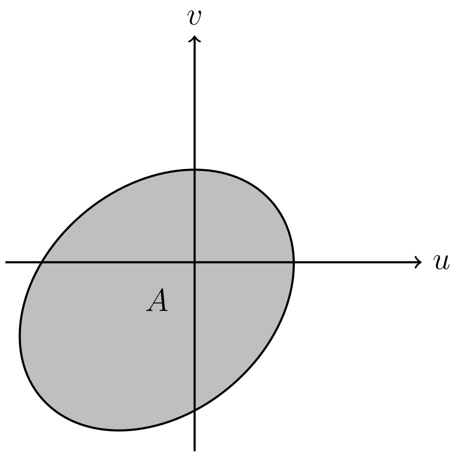
$$F_{X,Y}(x, y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f_{X,Y}(u, v)\, du dv,$$

så kallar vi $f_{X,Y}$ för variabelns **simultana täthetsfunktion**. Detta är typfallet för att (X, Y) är en kontinuerlig tvådimensionell stokastisk variabel.

Sannolikheten $F_{X,Y}(x, y)$ och sannolikheten för en mer generell delmängd $A \subset \mathbf{R}^2$ kan grafiskt illustreras genom figuren nedan. Observera att sannolikheten inte är den skuggade arean, utan volymen som uppstår när vi har en funktionsyta definierad ovanför det skuggade området. Arean symboliserar alltså ett integrations- eller summationsområde. Vi ”summerar” (via en integral eller summa) sedan sannolikhetsvärden för de intressanta värdena för variabeln.



Den halvoändliga rektangeln $C(x, y)$.



$P((X, Y) \in A)$ är sannolikheten att få ett resultat (x, y) som ligger i mängden A.

Egenskaper hos sannolikhetsfunktionen

(i) $p_{X,Y}(j, k) \geq 0$ för alla (j, k).

(ii) $\displaystyle\sum_j \sum_k p_{X,Y}(j, k) = 1$.

(iii) Om $A \subset \mathbf{R}^2$ så är $P(X \in A) = \displaystyle\sum\sum_{(j,k)\in A} p_{X,Y}(j, k)$.

Egenskaper hos den simultana täthetsfunktionen

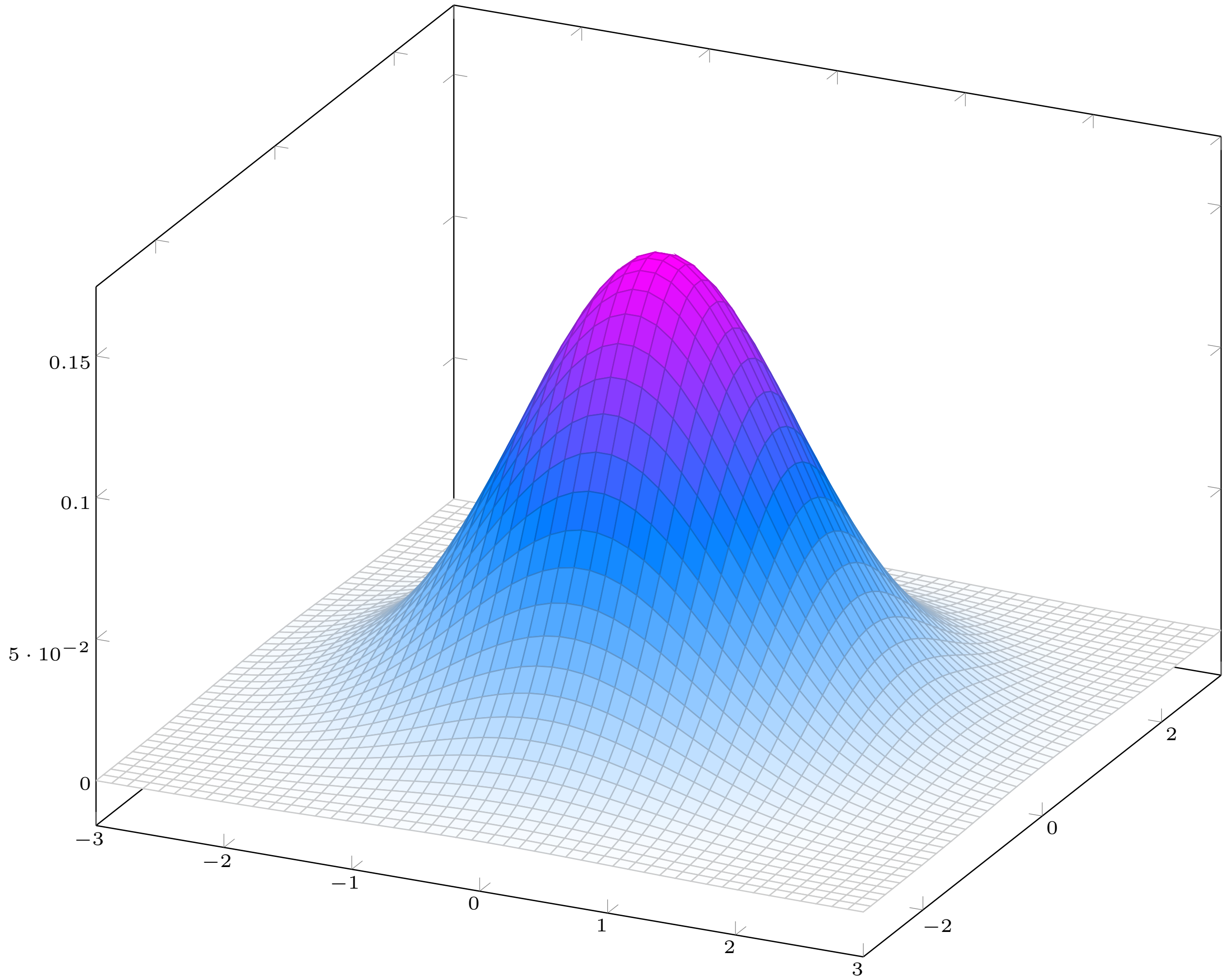
(i) $f_{X,Y}(x, y) \geq 0$ för alla $(x, y) \in \mathbf{R}^2$.

(ii) $\displaystyle\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, dxdy = 1$.

(iii) Om $A \subset \mathbf{R}^2$ är *snäll* så är $P((X, Y) \in A) = \displaystyle\iint_A f_{X,Y}(x, y)\, dxdy$.

(iv) Talet $f_{X,Y}(x, y)$ anger hur mycket sannolikhetsmassa det finns per areaenhet i punkten (x, y).

Exempel på hur en tvådimensionell täthetsfunktion kan se ut. Det är nu volymen, inte arean, som ska vara ett.



Exempel

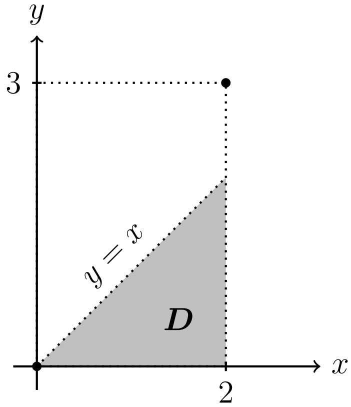
Det enklaste exemplet på en 2D-täthetsfunktion är den likformiga fördelningen. Låt A vara rektangeln med hörn i $(0,0)$ och $(2,3)$ och låt (X,Y) vara likformigt fördelad på A. Då ges $f_{X,Y}$ av $f_{X,Y}(x,y) = 1/6$ om $(x,y) \in A$ och $f_{X,Y}(x,y) = 0$ för övrigt. Hitta sannolikheten att $X > Y$.

Lösning. Det är alltid bra att försöka rita en figur, så vi gör detta parallellt med att vi betraktar arean av de delar vi är intresserad av.

Vi finner alltså denna sannolikhet genom att titta på hur stor del av A där detta villkor (dvs att $x > y$) är uppfyllt, vilket är triangeln med hörn i $(0,0)$, $(2,0)$ och $(2,2)$. Vi erhåller alltså

$$P(X > Y) = \frac{\text{area triangel}}{\text{area rektangel}} = \frac{2}{6} = \frac{1}{3}.$$

Testa att ställa upp en integral. Blir svaret detsamma?



Definition. De **marginella** täthetsfunktionerna f_X och f_Y för X och Y i en kontinuerlig stokastisk variabel (X, Y) ges av

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

Motsvarande gäller om (X, Y) är diskret:

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

Man kan även definiera marginella fördelningsfunktioner genom

$$F_X(x) = \lim_{y\to\infty} F_{X,Y}(x, y) \qquad \text{och} \qquad F_Y(y) = \lim_{x\to\infty} F_{X,Y}(x, y).$$

Så vad är egentligen dessa *marginella* funktioner? Det vi gör är att vi summerar alla möjligheter för den variabel vi inte är intresserade av och på det sättet skapar något som bara beror på en variabel. Detta leder också till följande sats.

Oberoende variabler

Sats. Om (X, Y) är en stokastisk variabel med simultan täthetsfunktion $f_{X,Y}$ gäller att X och Y är oberoende om och endast om $f_{X,Y}(x, y) = f_X(x)f_Y(y)$. För en diskret variabel är motsvarande villkor $p_{X,Y}(j, k) = p_X(j)p_Y(k)$.

Det är även sant att (i båda fallen) X och Y är oberoende om och endast om

$$F_{X,Y}(x, y) = F_X(x)F_Y(y).$$

Exempel

Låt (X, Y) ha sannolikhetsfunktionen $p_{X,Y}(j, k) = c(jk + k^3)$ för $j = 0, 1, 2, 3$ och $k = 0, 1$. Annars är $p_{X,Y} = 0$.

(i) Vad är c?

(ii) Bestäm $p_X(j)$ och $p_Y(k)$. Är X och Y oberoende?

(iii) Beräkna $P(X \leq 2, Y \geq 0.5)$.

Lösning:

(i) Vi måste välja $c > 0$ för att $p_{X,Y}$ ska kunna vara en sannolikhetsfunktion. För att finna c summerar vi över alla j och k:

$$\sum_{j=0}^{3}\sum_{k=0}^{1} c(jk + k^3) = \sum_{j=0}^{3} c(j + 1) = c(1 + 2 + 3 + 4) = 10c,$$

så $c = 1/10$ är nödvändigt och tillräckligt.

(ii) De marginella sannolikhetsfunktionerna fås ur definitionen enligt

$$p_X(j) = \sum_{k=0}^{1} c(jk + k^3) = c(j+1), \quad j = 0, 1, 2, 3.$$

$$p_Y(k) = \sum_{j=0}^{3} c(jk + k^3) = c(k^3 + k + k^3 + 2k + k^3 + 3k + k^3) = 2c(3k + 2k^3), \quad k = 0, 1.$$

Vi testar även att

$$\sum_{j=0}^{3} p_X(j) = c(1+2+3+4) = 1,$$
$$\sum_{k=0}^{1} p_Y(k) = 2c(3+2) = 1,$$

så p_X och p_Y är sannolikhetsfunktioner. Vi undersöker oberoendet:

$$p_X(j)p_Y(k) = 2c^2(j+1)(3k+2k^3).$$

Här kan man kanske direkt tro att X och Y är beroende, men det skulle vara en gissning. Vi undersöker explicit:

$p_{X,Y}$	$j=0$	$j=1$	$j=2$	$j=3$
$k=0$	0	0	0	0
$k=1$	1/10	2/10	3/10	4/10

$p_X p_Y$	$j=0$	$j=1$	$j=2$	$j=3$
$k=0$	0	0	0	0
$k=1$	$\frac{1\cdot 2\cdot 5}{100}$	$\frac{2\cdot 2\cdot 5}{100}$	$\frac{3\cdot 2\cdot 5}{100}$	$\frac{4\cdot 2\cdot 5}{100}$

Vi ser här att alla siffror matchar varandra och att därmed $p_{X,Y}(j,k) = p_X(j)p_X(k)$ för *alla* j och k. Detta trots att uttrycken såg väldigt olika ut från början. Var försiktiga med att dra slutsatser utan att undersöka ordentligt!

(iii) Vi kan numrera de tillåtna värdena (där $X = j \leq 2$ och $Y = k \geq 0.5$) på (j,k): $(0,1)$, $(1,1)$, $(2,1)$. Sannolikheten blir då

$$P(X \leq 2, Y \geq 0.5) = p_{X,Y}(0,1) + p_{X,Y}(1,1) + p_{X,Y}(2,1) = c(1+2+3) = 6/10.$$

Exempel

Låt (X,Y) ha den simultana täthetsfunktionen

$$f_{X,Y}(x,y) = \begin{cases} c(x^2y + xy^2), & 0 \leq x \leq 1,\ 0 \leq y \leq 1, \\ 0, & \text{annars.} \end{cases}$$

Lös följande problem.

(i) Bestäm c;

(ii) Vad är sannolikheten för händelsen att $X > 1/2$ och att $Y < 1/2$? Det vill säga, beräkna sannolikheten $P(X > 1/2,\ Y < 1/2)$;

Lösning:

(i) Vi räknar ut följande

$$1 = \iint_{\mathbf{R}^2} f_{X,Y}(x,y)\,dxdy = \int_0^1 \int_0^1 c(x^2y + xy^2)\,dxdy = c\int_0^1 \left(\frac{y}{3} + \frac{y^2}{2}\right)\,dy = \frac{c}{3},$$

så $c = 3$ är nödvändigt för att få en täthetsfunktion.

(ii) Vi har

$$P(X > 1/2 \text{ och } Y < 1/2) = \int_{1/2}^1 \int_0^{1/2} 3(x^2y + xy^2)\,dydx = \frac{5}{32}.$$

Exempel

Med samma täthetsfunktion som i föregående exempel, lös följande problem.

(i) Beräkna $P(X > 1/2)$;

(ii) Beräkna $P\big(Y < 1/2 \mid X > 1/2\big)$;

(iii) Beräkna $f_X(x)$ och $f_Y(y)$. Är X och Y oberoende?

(iv) Beräkna $F_{X,Y}(x,y)$.

(v) Vad blir $\dfrac{\partial^2}{\partial_x \partial_y} F_{X,Y}(x,y)$?

Lösning:

(i) Vi har endast ett krav på x, så vi integrerar över alla y:

$$P(X > 1/2) = \int_{1/2}^1 \int_0^1 3(x^2y + xy^2)\,dydx = \frac{13}{16}.$$

(ii) Enligt definitionen av betingad sannolikhet,

$$P\big(Y < 1/2 \mid X > 1/2\big) = \frac{P(X > 1/2,\ Y < 1/2)}{P(X > 1/2)} = \frac{5/32}{13/16} = \frac{5}{26}.$$

(iii) De marginella tätheterna beräknas direkt från definitionen:

$$f_X(x) = \int_0^1 3(xy^2 + x^2y)\,dy = x + \frac{3}{2}x^2, \quad 0 \le x \le 1,$$
$$f_Y(y) = \int_0^1 3(xy^2 + x^2y)\,dx = y + \frac{3}{2}y^2, \quad 0 \le y \le 1.$$

Vi ser tydligt att $f_X(x)f_Y(y) \neq f_{X,Y}(x,y)$ för många val av punkter (x,y); ta till exempel $(x,y) = (1,1)$. Variablerna är beroende.

(iv) Låt $(x, y) \in [0,1] \times [0,1]$. Då blir

$$F_{X,Y}(x,y) = \int_0^x \int_0^y 3(uv^2 + u^2 v)\,dv du = \frac{x^3y^2 + x^2y^3}{2}.$$

Om $x < 0$ *eller* $y < 0$, måste $F_{X,Y}(x,y) = 0$, och om $x > 1$ *och* $y > 1$, så måste $F_{X,Y}(x,y) = 1$. Övriga fall täcks av

$$F_{X,Y}(x,y) = \frac{y^2 + y^3}{2}, \quad 0 \leq y \leq 1 \text{ och } x > 1,$$
$$F_{X,Y}(x,y) = \frac{x^3 + x^2}{2}, \quad 0 \leq x \leq 1 \text{ och } y > 1.$$

(v) Om $0 \leq x \leq 1$ och $0 \leq y \leq 1$,

$$\frac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = \frac{\partial}{\partial x}\left(\frac{2x^3y + 3x^2y^2}{2}\right) = \frac{6x^2y + 6xy^2}{2} = f_{X,Y}(x,y).$$

Övriga kombinationer av x och y kommer att ge $\dfrac{\partial^2}{\partial x \partial y} F_{X,Y}(x,y) = 0$.

3 Väntevärden

Väntevärden av funktioner av flerdimensionella stokastiska variabler fungerar analogt med tidigare. Låt oss formulera en sats.

Väntevärde och funktioner av stokastiska variabler

Sats. Låt $Y = g(X)$ och $W = h(U,V)$. I de kontinuerliga fallen blir

$$E(Y) = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx \qquad \text{och} \qquad E(W) = \iint_{\mathbf{R}^2} h(x,y) f_{U,V}(x,y)\,dxdy,$$

och om X, U, V är diskreta:

$$E(Y) = \sum_k g(k) p_X(k) \qquad \text{och} \qquad E(W) = \sum_j \sum_k h(j,k) p_{U,V}(j,k).$$

Det kontinuerliga fallet (med täthetsfunktion) kommer vi inte åt på något annat sätt än att egentligen ta satsen ovan som definition. Om vi tänker på punkt (ii) i rutan föregående satsen, så förefaller det ganska rimligt. Sammansättningen $g(Y)$ till exempel är en ny stokastisk variabel och då skulle

$$E(g(Y)) = \int_\Omega g(Y)(\omega)\,dP(\omega),$$

vilket är precis hur satsen tolkas. I det diskreta fallet kan vi faktiskt producera ett bevis:

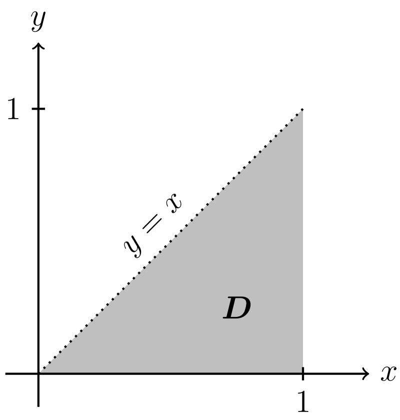
$$\begin{aligned} E(g(X)) &= \sum_k kP(g(X) = k) = \sum_k k \sum_{m\,:\,g(m)=k} P(X = m) \\ &= \sum_m \sum_{k\,:\,g(m)=k} kP(X = m) = \sum_m P(X = m) \sum_{k\,:\,g(m)=k} k = \sum_m g(m) P(X = m), \end{aligned}$$

om vi antar att serien är absolutkonvergent i det oändliga fallet så vi kan byta summationsordningen. Rent praktiskt kan beräkningarna gå till på följande sätt.

Exempel

Låt $f_{X,Y}(x,y) = 2$ om $0 < y < x < 1$ och $f_{X,Y}(x,y) = 0$ för övrigt. Bestäm väntevärdet $E(XY + Y^2X)$.

Lösning: Vi använder satsen ovan:



$$
\begin{aligned}
E(XY + Y^2X) &= \iint_{\mathbf{R}^2} (xy + y^2x) f_{X,Y}(x,y)\,dxdy \\
&= 2\int_0^1 \int_0^x (xy + y^2x)\,dydx \\
&= 2\int_0^1 \left[\frac{xy^2}{2} + \frac{xy^3}{3}\right]_0^x dx \\
&= \int_0^1 \left(x^3 + \frac{2x^4}{3}\right) dx \\
&= \frac{1}{4} + \frac{2}{15} = \frac{23}{60}.
\end{aligned}
$$

Exempel

Låt X och Y vara oberoende med $E(X) = 2$, $E(X^2) = 8$, $E(Y) = -1$, $V(Y) = 2$. Beräkna $E(XY)$, $E(X^2Y^2)$ och $D(2X - 3Y)$.

Lösning. Här behöver vi inte använda satsen direkt. Eftersom variablerna är oberoende blir $E(XY) = E(X)E(Y) = 2 \cdot (-1) = -2$, och

$$E(X^2Y^2) = E(X^2)E(Y^2) = 8E(Y^2) = 8(V(Y) + E(Y)^2) = 8 \cdot 3 = 24.$$

Vidare erhåller vi

$$V(2X - 3Y) = 2^2V(X) + (-3)^2V(Y) = 4(8 - 2^2) + 9 \cdot 2 = 34,$$

så $D(2X - 3Y) = \sqrt{34}$.

Exempel

Låt X och Y vara oberoende likafördelade variabler med $p_X(k) = p_Y(k) = 1/4$ om $k = 1, 2$ och $p_X(k) = p_Y(k) = 1/2$ om $k = 0$. Bestäm sannolikhetsfuntionen för $Z = X + Y$, väntevärdet $E(X + Y)$ samt variansen $V(X + Y)$.

Lösning. Vi börjar med att bestämma $p_Z(k)$.

$Z = X + Y = n$	Par (j,k) så $j + k = n$	Total sannolikhet
-1	–	0
0	$(0,0)$	$\frac{1}{2}\frac{1}{2} = \frac{1}{4}$
1	$(0,1)$, $(1,0)$	$2 \cdot \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{4}$
2	$(0,2)$, $(1,1)$, $(2,0)$	$2 \cdot \frac{1}{2}\frac{1}{4} + \frac{1}{4}\frac{1}{4} = \frac{5}{16}$
3	$(1,2)$, $(2,1)$	$2 \cdot \frac{1}{4}\frac{1}{4} = \frac{1}{8}$
4	$(2,2)$	$\frac{1}{4}\frac{1}{4} = \frac{1}{16}$
5	–	0

Två sätt att beräkna väntevärdet:

$$E(Z) = E(X+Y) = E(X) + E(Y) = 2E(X) = 2\sum_{k=0}^{2} kp_X(k) = 2\left(0 + \frac{1}{4} + \frac{2}{4}\right) = \frac{3}{2},$$

$$E(Z) = \sum_{k=0}^{4} kp_Z(k) = \frac{1}{4} + \frac{10}{16} + \frac{3}{8} + \frac{4}{16} = \frac{24}{16} = \frac{3}{2}.$$

Möjligen kan man tycka att det andra alternativet ser enklare ut, men då har man glömt bort allt arbete som gick åt till att skapa tabellen ovan. Det första alternativet är nästan alltid det enklaste. Variansen beräknar man enklast enligt följande

$$V(Z) = V(X+Y) = \big[\text{oberoende variabler}\big] = V(X) + V(Y) = 2V(X) = \frac{1}{8}$$

eftersom vi vet att

$$V(X) = E(X^2) - E(X)^2 = \sum_{k=0}^{2} k^2 p_X(k) - \frac{9}{16} = \left(0 + \frac{1}{4} + \frac{4}{4}\right) - \frac{9}{16} = \frac{1}{16}.$$

4 Kovarians

Vi har betraktat variansen för stokastiska variabler, men kan man säga någon om variationen mellan två variabler utan att stirra sig blind på sannolikhets- eller täthetsfunktionen? Visst kan man det, och svaret kommer i form av kovarians eller korrelation

Definition. Låt X och Y vara två stokastiska variabler sådana att $E(X) = \mu_X$, $V(X) = \sigma_X^2$, $E(Y) = \mu_Y$ samt $V(Y) = \sigma_Y^2$. Kovariansen $C(X,Y)$ definieras enligt

$$C(X,Y) = E\big((X-\mu_X)(Y-\mu_Y)\big)$$

och korrelationen mellan X och Y enligt

$$\rho(X,Y) = \frac{C(X,Y)}{\sigma_X\,\sigma_Y}.$$

Både kovarians och korrelation är ett mått på linjärt beroende mellan X och Y där korrelationen är normerad så det går att jämföra olika fall.

Okorrelerad

Definition. Om $C(X,Y) = 0$ kallas X och Y för okorrelerade.

Kovariansen uppfyller följande egenskaper.

(i) $C(X,Y) = E(XY) - E(X)E(Y)$.

(ii) Om X och Y är oberoende så är $C(X,Y) = 0$.

(iii) $|\rho(X,Y)| \leq 1$ med likhet om och endast om det finns ett linjärt samband mellan X och Y.

(iv) $C(X,X) = V(X)$.

(v) $C\left(a_0 + \sum_{i=1}^{m} a_i X_i,\ b_0 + \sum_{j=1}^{n} b_j Y_j\right) = \sum_{i=1}^{m}\sum_{j=1}^{n} a_i b_j C(X_i, Y_j)$.

Bevis.

(i) Detta följer från i princip samma argument som Steiners sats. Låt oss expandera definition av kovarians:

$$C(X,Y) = E\big((X-\mu_X)(Y-\mu_Y)\big) = E(XY) - \mu_X E(Y) - \mu_Y E(X) + \mu_X \mu_Y = E(XY) - \mu_X \mu_Y,$$

vilket är precis vad vi ville visa.

(ii) Eftersom X och Y är oberoende, så följer det att $E(XY) = E(X)E(Y)$. Detta leder givetvis med föregående punkt i tanken till att $C(X,Y) = 0$.

(iii) Väntevärdet av en kvadrat är givetvis icke-negativt, så

$$\begin{aligned} 0 &\leq E\left(\left(\frac{X-\mu_X}{\sigma_X} \pm \frac{Y-\mu_Y}{\sigma_Y}\right)^2\right) \\ &= \frac{1}{\sigma_X^2} E((X-\mu_X)^2) + \frac{1}{\sigma_Y^2} E((Y-\mu_Y)^2) \pm 2\frac{E((X-\mu_X)(Y-\mu_Y))}{\sigma_X \sigma_Y} \\ &= 2\left(1 \pm \rho(X,Y)\right). \end{aligned}$$

Alltså måste $|\rho(X,Y)| \leq 1$. Dessutom ser vi på köpet att om $\rho(X,Y) = \pm 1$ så måste väntevärdet av kvadraten vara lika med noll, så

$$\frac{X-\mu_X}{\sigma_X} \pm \frac{Y-\mu_Y}{\sigma_Y} = 0 \quad \Leftrightarrow \quad Y = \mu_Y \pm \frac{\sigma_Y(X-\mu_X)}{\sigma_X} \quad \Leftrightarrow \quad Y = aX + b.$$

Således är beroendet mellan X och Y i form av en rät linje. En lite kvalifikation är nödvändig: det är en integral som blir noll och vi drar slutsatsen att integranden är noll. Detta är inte helt självklart. Nu råkar vi veta att integranden är icke-negativ, så vi har ingen negativ area som spökar. Men vi kan fortfarande modifiera integranden här och där utan att ändra integralen, så viss försiktighet är nödvändig. Formellt heter detta att likheten gäller *nästan överallt*. Men vi är tillbaka till Lebesgue-integralen nu, så låt oss lämna området kvickt.

(iv) Vi ser direkt att

$$C(X,X) = E(X^2) - E(X)^2 = V(X)$$

vilket är variansen enligt Steiners sats.

(v) Ohyggligt åbäke som lämnas som övning.. skämt åsido så kommer ni stöta på detta igen, men då med metoder från linjär algebra i bagaget så saker och ting faktiskt går att hantera utan tårar.

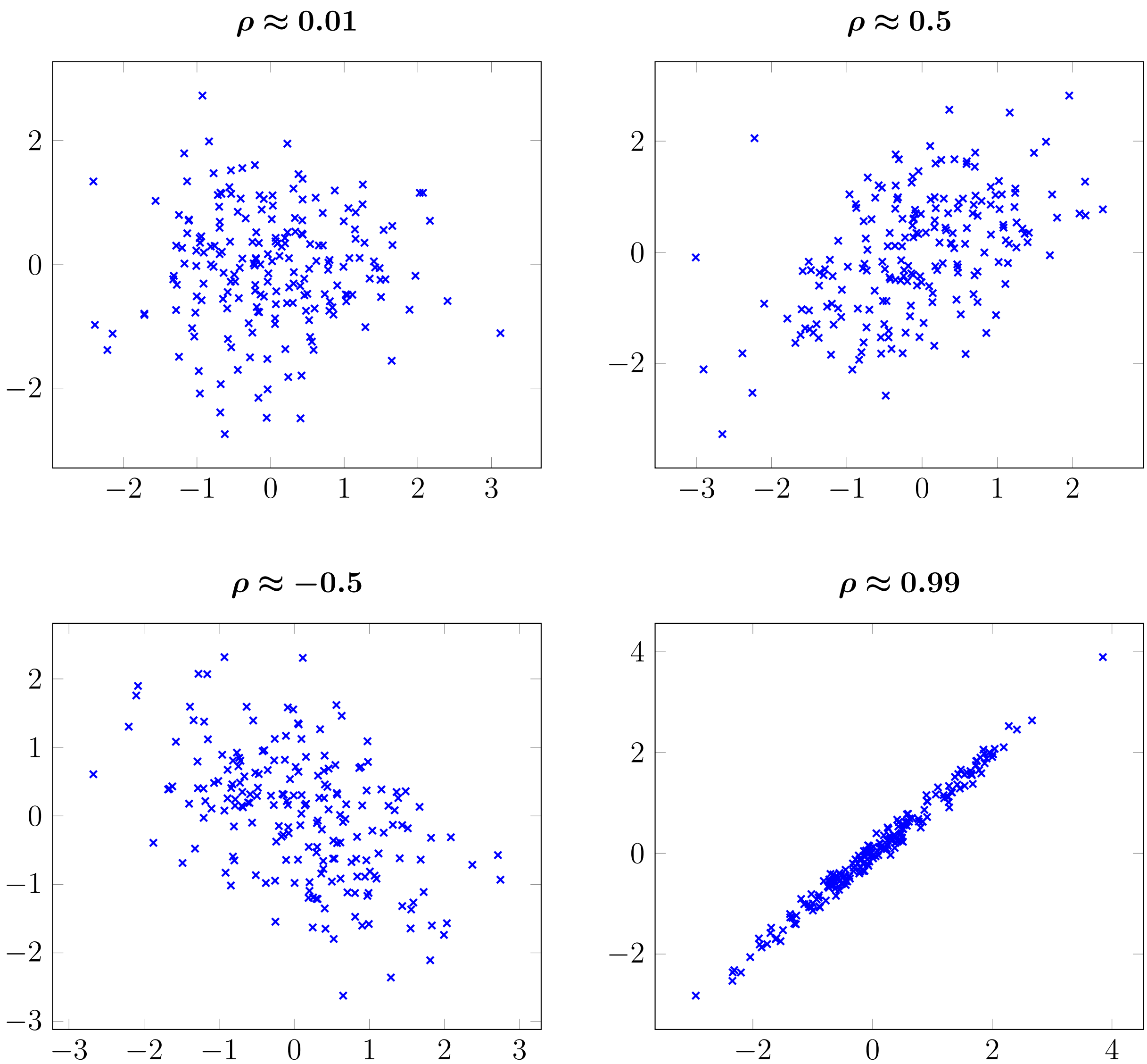
Observera att $C(X,Y) = 0$ *inte* nödvändigtvis innebär oberoende. Låt till exempel X vara rektangelfördelad enligt $X \sim \text{Re}(-1,1)$ och definiera $Y = X^2$. Uppenbarligen beroende variabler, men

$$C(X,Y) = E(XY) - E(X)E(Y) = E(X^3) - 0 \cdot E(Y) = E(X^3) = \int_{-1}^{1} x^3 \cdot \frac{1}{2}\,dx = 0,$$

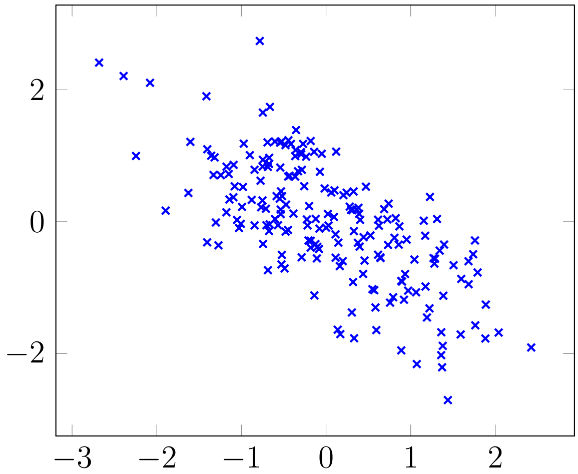
så X och Y är okorrelerade.

5 Vad innebär korrelationen grafiskt?

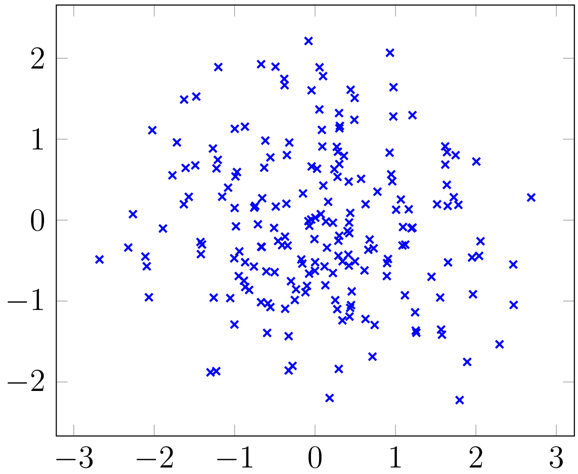
Korrelation används ofta för att avgöra om det finns ett linjärt samband mellan x och y-värden när vi på något sätt fått en samling data $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$. Även detta återkommer i nästa del av kursen!



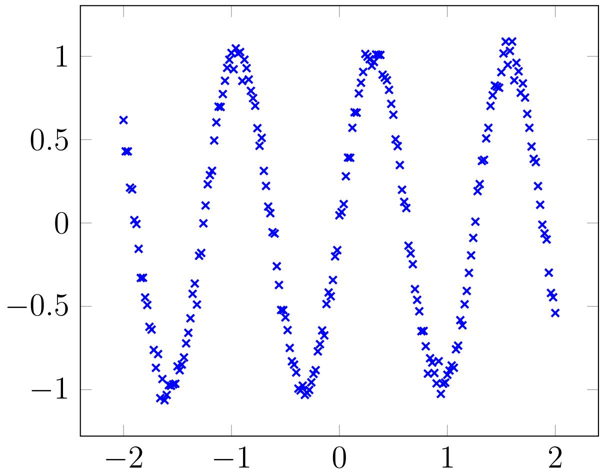
$\rho \approx -0.7$



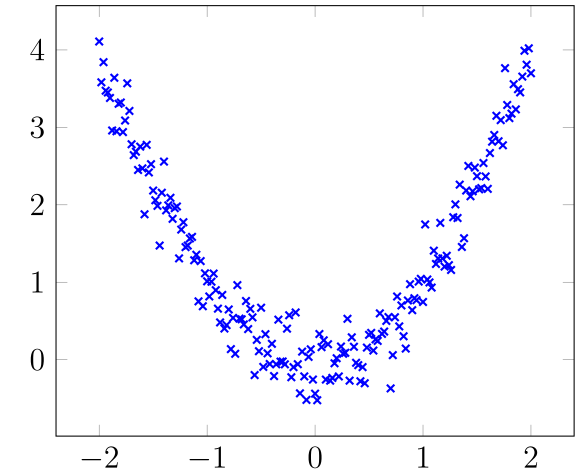
$\rho \approx -0.1$



$\rho \approx 0.2$ (wtf?)



$\rho \approx 0$



6 Funktioner av stokastiska variabler

Exempel

Om $X \sim \text{Exp}(\lambda)$, vad får $Y = e^X$ för täthetsfunktion?

Lösning: Vad blir f_Y? Vi ser att $Y > 0$ från definitionen så $f_Y(y) = 0$ för $y \leq 0$. Vi ställer upp $F_Y(y)$ för $y > 0$:

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y) = P(X \leq \log y) = F_X(\log y), \quad y > 0.$$

Vi deriverar fram $f_Y(y) = F_Y'(y) = \frac{1}{y} \cdot f_X(\log y)$. Vi vet att $f_X(x) = \lambda e^{-\lambda x}$ för $x > 0$ och om $x \leq 0$ blir $f_X(x) = 0$. Eftersom $\log y < 0$ då $0 < y < 1$ får vi två fall:

$$f_Y(y) = \begin{cases} \frac{\lambda}{y} e^{-\lambda \log y}, & y > 1, \\ 0, & y \leq 1, \end{cases} = \begin{cases} \lambda y^{-1-\lambda}, & y > 1, \\ 0, & y \leq 1. \end{cases}$$

Exempel

Om $X \sim \text{Re}(-1,2)$, vad får $Y = X^2$ för täthetsfunktion?

Lösning: Om $y < 0$ så måste $F_Y(y) = P(Y \leq y) = 0$ ($Y = X^2$ kommer aldrig att vara negativ). Låt oss anta att $y \geq 0$. Vi beräknar

$$\begin{aligned} F_Y(y) &= P(Y \leq y) = P(X^2 \leq y) = P(-\sqrt{y} \leq X \leq \sqrt{y}) \\ &= P(X \leq \sqrt{y}) - P(X \leq -\sqrt{y}) = F_X(\sqrt{y}) - F_X(-\sqrt{y}). \end{aligned}$$

Vi antar att f_Y är kontinuerlig och deriverar fram ett uttryck:

$$f_Y(y) = F_Y'(y) = \begin{cases} \frac{1}{2\sqrt{y}}\big(f_X(\sqrt{y}) + f_X(-\sqrt{y})\big), & y > 0, \\ 0, & y \leq 0. \end{cases}$$

Vidare vet vi att $f_X(x) = 1/3$ om $-1 \leq x \leq 2$ och $f_X(x) = 0$ annars, så

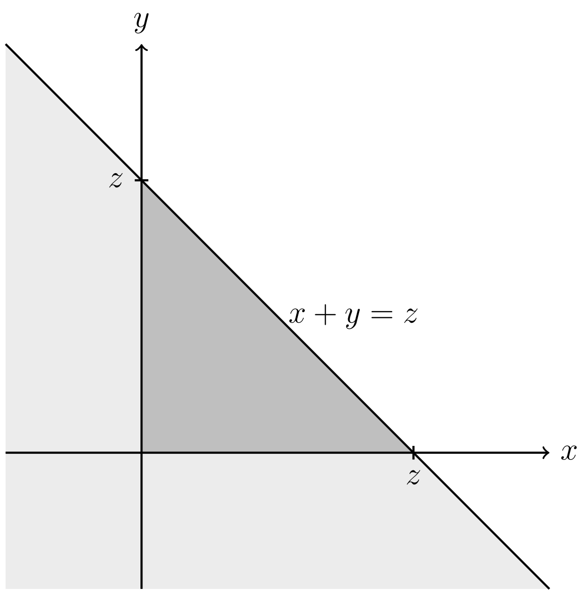
$$f_Y(y) = \begin{cases} 0, & y < 0, \\ \frac{1}{3\sqrt{y}}, & 0 \leq y < 1, \\ \frac{1}{6\sqrt{y}}, & 1 \leq y < 4, \\ 0, & y \geq 4. \end{cases}$$

Om vi har den simultana täthets- eller sannolikhetsfunktionen så kan vi hitta fördelningar för funktioner av flera stokastiska variabler. Låt oss betrakta ett par vanliga exempel.

Exempel

Låt $X \sim \text{Exp}(1)$ och $Y \sim \text{Exp}(2)$ vara oberoende. Beräkna täthetsfunktionen för $Z = X+Y$.

Klart att $f_{X,Y}(x,y) = f_X(x)f_Y(y) = 2 \cdot 1 \cdot e^{-x-2y}$ för $x \geq 0$ och $y \geq 0$ (annars $f_{X,Y} = 0$). Vi söker $f_Z(z)$. Det är klart att om $z < 0$ så är $f_{X,Y}(x,y) = 0$. Antag att $z \geq 0$. Vi ställer upp fördelningsfunktionen $F_Z(z)$ för Z, och för det behöver vi ha klart för oss vilket område vi integrerar över:



Fördelningsfunktionen blir nu, för $z \geq 0$,

$$
\begin{aligned}
F_Z(z) = P(X+Y \leq z) &= \int_{-\infty}^{z} \int_{-\infty}^{z-x} f_{X,Y}(x,y)\,dydx \\
&= 2\int_0^z e^{-x} \int_0^{z-x} e^{-2y}\,dydx = \int_0^z e^{-x}(1-e^{-2(z-x)})\,dx \\
&= \int_0^z \left(e^{-x} - e^{-2z+x}\right) dx = 1 + e^{-2z} - 2e^{-z},
\end{aligned}
$$

och vi kan derivera fram $f_Z(z) = F_Z'(z) = 2e^{-z} - 2e^{-2z}$. Kontrollera att $f_Z(z) \geq 0$ samt att $\int_0^{\infty} f_Z(z)\,dz = 1$.

Faltningssatsen

Sats. Om X och Y är oberoende kontinuerliga stokastiska variabler så ges täthetsfunktionen f_Z för $Z = X + Y$ av

$$
f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,dx, \quad z \in \mathbf{R}.
$$

Motsvarande gäller för diskreta variabler:

$$
p_Z(k) = \sum_j p_X(j) p_Y(k-j).
$$

Min och max

Vad blir fördelningsfunktionerna för maximum respektive minimum av två oberoende stokastiska variabler?

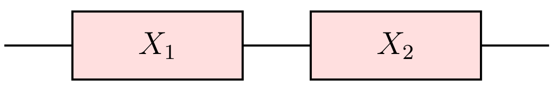
Låt $Y = \max(X_1, X_2)$. Då blir

$$
\begin{aligned}
F_Y(y) &= P(\max(X_1,X_2) \leq y) = P(X_1 \leq y \text{ och } X_2 \leq y) = P(X_1 \leq y)P(X_2 \leq y) \\
&= F_{X_1}(y) F_{X_2}(y).
\end{aligned}
$$

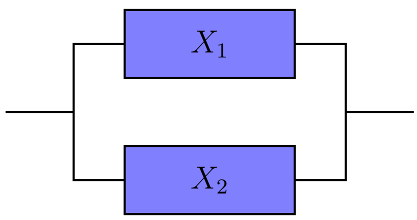
För $Y = \min(X_1, X_2)$ så erhåller vi

$$
\begin{aligned}
F_Y(y) &= P(\min(X_1,X_2) \leq y) = 1 - P(\min(X_1,X_2) > y) = 1 - P(X_1 > y \text{ och } X_2 > y) \\
&= 1 - P(X_1 > y)P(X_2 > y) = 1 - (1 - P(X_1 \leq y))(1 - P(X_2 \leq y)) \\
&= 1 - (1 - F_{X_1}(y))(1 - F_{X_2}(y)).
\end{aligned}
$$

En vanlig situation där dessa uttryck dyker upp är i parallell- och seriekoppling av system. Schematiskt kan man beskriva dessa situationer med blockscheman.



Seriekoppling. Båda kanalerna måste fungera. Ger minimum av livslängderna.



Parallellkoppling. Räcker att en kanal fungerar. Ger maximum av livslängderna.

7 ($\star\star$)Formell definition av 2D-stokastisk variabel

I $\mathbf{R}^2$ är Borelfamiljen $\mathscr{B}$ den minsta σ-algebra som innehåller alla öppna rektanglar $(a, b)\times(c, d)$.

Stokastisk variabel

Definition. En tvådimensionell stokastisk variabel är en reell-vektorvärd funktion (X, Y) (två komponenter) definierad på ett utfallsrum Ω. Alltså avbildar (X, Y) olika utfall på reella vektorer; $(X, Y)\colon \Omega \to \mathbf{R}^2$. Vi kräver att $(X, Y)^{-1}(B) \in \mathcal{F}$ för alla $B \in \mathscr{B}$. Algebran $\mathcal{F}$ är mängden av alla tillåtna händelser. Om (X, Y) bara antar ändligt eller uppräkneligt många värden så kallar vi (X, Y) för en diskret stokastisk variabel. Om varken X eller Y är diskret kallar vi (X, Y) för kontinuerlig.

8 ($\star$)Bökigt exempel

Två sätt att räkna ut $E(g(X))$

Låt $\Theta \sim \mathrm{Re}(0, \pi/4)$ vara likformigt fördelad och definiera $Y = \cos\Theta$. Vad blir $E(Y)$?

Lösning: Det enklaste sättet är att använda satsen från början av föreläsningen:

$$E(Y) = E(\cos\Theta) = \int_{-\infty}^{\infty} \cos\theta f_\Theta(\theta)\, d\theta = \frac{4}{\pi}\int_0^{\pi/4} \cos\theta\, d\theta = \frac{4}{\pi}\left(\frac{\sqrt{2}}{2} - 0\right) = \frac{2\sqrt{2}}{\pi}.$$

Den andra varianten börjar med beräkning av täthetsfunktionen för $Y = \cos\Theta$. Vi ställer upp fördelningsfunktionen först:

$$F_Y(y) = P(Y \le y) = P(\cos\Theta \le y) = P(\Theta \ge \arccos y) = 1-P(\Theta < \arccos y) = 1-F_\Theta(\arccos y).$$

Här har vi utnyttjat att $\cos\theta$ är avtagande för $\theta \in [0, \pi/4]$. Vi kan nu derivera fram $f_Y(y)$:

$$\begin{aligned} f_Y(y) = F'_Y(y) &= -F'_\Theta(\arccos y)\cdot\frac{-1}{\sqrt{1-y^2}} = \frac{f_\Theta(\arccos y)}{\sqrt{1-y^2}} \\ &= \begin{cases} \frac{4}{\pi}\frac{1}{\sqrt{1-y^2}}, & 0 \le \arccos y \le \frac{\pi}{4} \quad\Leftrightarrow\quad \frac{\sqrt{2}}{2} \le y \le 1 \\ 0, & \text{för övrigt.} \end{cases} \end{aligned}$$

Vi kan nu beräkna $E(Y)$ enligt definitionen:

$$E(Y) = \int_{-\infty}^{\infty} y f_Y(y)\, dy = \frac{4}{\pi}\int_0^{\sqrt{2}/2} \frac{y}{\sqrt{1-y^2}}\, dy = \frac{4}{\pi}\left[-\sqrt{1-y^2}\right]_0^{\frac{\sqrt{2}}{2}} = \frac{2\sqrt{2}}{\pi}.$$

Vilket metod tycker du är enklast?

9 ($\star\star$)Normalfördelning

Täthetsfunktionen φ föreslagen för normalfördelningen $N(0, 1)$ är en täthetsfunktion. Beviset är så roligt att jag inte kan låta bli, men det kräver lite flervariabelanalys. Vi behöver visa att

$$\int_{-\infty}^{\infty} e^{-x^2/2}\, dx = \sqrt{2\pi}.$$

Sättet vi kommer åt det hela är genom att titta på kvadraten och tolka denna som en itererad dubbelintegral:

$$\begin{aligned}\left(\int_{-\infty}^{\infty} e^{-x^2/2}\,dx\right)^2 &= \left(\int_{-\infty}^{\infty} e^{-x^2/2}\,dx\right)\left(\int_{-\infty}^{\infty} e^{-y^2/2}\,dy\right) = \iint_{\mathbf{R}^2} e^{-(x^2+y^2)/2}\,dx\,dy \\ &= \int_0^{\infty}\int_0^{2\pi} re^{-r^2/2}\,d\theta\,dr = 2\pi\left[-e^{-r^2/2}\right]_0^{\infty} = 2\pi.\end{aligned}$$