

Flerdimensionella fördelningar

19 okt. 2026, 9 min läsning

UTKAST

#matematik #sannolikhetslära #grundbegrepp

Kurs: S0008M Förkunskaper: Fördelningsfunktion, Funktioner av flera variabler

En flerdimensionell fördelning beskriver två eller flera stokastiska variabler samtidigt genom en simultan fördelning. Ur den får vi marginalfördelningarna för varje variabel för sig, men marginalerna räcker inte för att återskapa simultanfördelningen om variablerna är beroende.

1. Tvådimensionella stokastiska variabler

En **tvådimensionell stokastisk variabel** är ett par (ξ, η) av **stokastiska variabler** som betraktas tillsammans — en punkt i planet vars koordinater är slumpmässiga. Ibland skriver vi (ξ_1, ξ_2) i stället för (ξ, η) .

Vi behöver den när variablerna **inte** kan antas vara oberoende: utsänd och mottagen signal, längd och vikt hos samma person, två studenters betyg på samma tenta. Är variablerna oberoende räcker det att känna fördelningarna var för sig; är de beroende måste vi beskriva dem tillsammans.

Som i det endimensionella fallet skiljer vi mellan diskreta och kontinuerliga fördelningar.

2. Diskret fall: simultan sannolikhetsfunktion

(ξ, η) är diskret om både ξ och η är diskreta. Fördelningen anges med den **simultana sannolikhetsfunktionen**

$$p(x, y) = P(\xi = x, \eta = y)$$

för de värden (x, y) som paret kan anta, och $p(x, y) = 0$ annars. Den skrivs ofta som en tabell med x i raderna och y i kolumnerna.

Sannolikheten att paret hamnar i ett område C i planet är summan av massan där:

$$P((\xi, \eta) \in C) = \sum_{(x,y) \in C} p(x,y)$$

3. Kontinuerligt fall: simultan frekvensfunktion

Enligt kompendiets definition 1.3 har (ξ, η) en **kontinuerlig tvådimensionell fördelning** om det finns en funktion $f(x, y)$ sådan att för varje område D i planet

$$P((\xi, \eta) \in D) = \iint_D f(x, y) dx dy$$

$f(x, y)$ kallas den **simultana frekvensfunktionen** och uppfyller

$$f(x, y) \geq 0, \quad \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x, y) dx dy = 1$$

Sannolikhet är alltså volym under ytan $z = f(x, y)$, precis som den i en dimension är area under **frekvensfunktionen**. Beräkningarna är **dubbelintegraler**.

Simultan fördelningsfunktion

Oavsett typ bestäms fördelningen av $F(x, y) = P(\xi \leq x, \eta \leq y)$. I det kontinuerliga fallet är $f(x, y) = \frac{\partial^2 F}{\partial x \partial y}$. Vi använder den sällan direkt, men idén att först ta fram en fördelningsfunktion och sedan derivera återkommer i **Icke-linjära funktioner av stokastiska variabler**.

4. Väntevärden av funktioner

Är $g(x, y)$ en funktion av två variabler gäller

$$E(g(\xi, \eta)) = \sum_x \sum_y g(x, y) p(x, y) \quad \text{respektive} \quad E(g(\xi, \eta)) = \iint_{\mathbb{R}^2} g(x, y) f(x, y) dx dy$$

Med $g(x, y) = x + y$ följer till exempel $E(\xi + \eta) = E(\xi) + E(\eta)$ direkt, och med $g(x, y) = xy$ får vi $E(\xi\eta)$ som behövs för **kovariansen**.

5. Likformig tvådimensionell fördelning

(ξ, η) är **likformigt fördelad** på ett område K om frekvensfunktionen är konstant på K och noll utanför (kompendiet, definition 1.4):

$$f(x, y) = \begin{cases} \frac{1}{|K|}, & (x, y) \in K \\ 0, & \text{annars} \end{cases}$$

där $|K|$ är arean av K . För ett delområde $A \subset K$ blir sannolikheten ett rent areaförhållande:

$$P((\xi, \eta) \in A) = \frac{|A|}{|K|}$$

Rita området

Med likformig fördelning blir nästan varje uppgift geometri: rita K , rita området som händelsen motsvarar, och räkna areor. Ingen integral behövs.

6. Oberoende

ξ och η är oberoende om och endast om den simultana fördelningen är produkten av **marginalfördelningarna**:

$$p(x, y) = p_\xi(x) p_\eta(y) \quad \text{respektive} \quad f(x, y) = f_\xi(x) f_\eta(y) \quad \text{för alla } x, y$$

Räcker det med ett enda par (x, y) där likheten inte gäller är variablerna beroende. Omvänt: är ξ och η oberoende och kända var för sig får vi den simultana fördelningen som produkten.

Vanliga fel

- **$f(x, y)$ som sannolikhet.** Frekvensfunktionen kan vara större än 1; sannolikhet är *volym* under ytan, inte funktionsvärdet.
- **Integrerat över hela kvadraten.** f är 0 utanför området, så integrationsgränserna följer området där $f > 0$.
- **Oberoende på ett icke-rektangulärt område.** Är området där $f > 0$ inte en rektangel, t.ex. triangeln $0 < y < x < 1$, är variablerna beroende – även om f är konstant och formeln ser ut att faktorisera.
- **Glömt $|K|$.** Likformig fördelning har höjden $1/|K|$, inte 1.

7. Lösta lektionsuppgifter

Övningarna märkta *Kompendiet* kommer ur kompendiet om flerdimensionella fördelningar och rekommenderas inför tentan. Kompendiet har inget facit, så varje sådan övning är löst två gånger oberoende av varandra.

≡ Lektion 17, övning 3 – Två oberoende rektangelfördelade variabler

Antag att ξ och η är oberoende och rektangelfördelade på intervallet $(0, 1)$.

- Bestäm den simultana frekvensfunktionen för (ξ, η) . Vad kallas denna tvådimensionella fördelning?
- Händelsen $A = \{\xi + \eta \leq 1,5\}$ kan skrivas $\{(\xi, \eta) \in C\}$ för ett lämpligt område C i planet. Vilket? Rita en bild.
- Beräkna $P(\xi + \eta \leq 1,5)$.
- Beräkna $P(\xi + \eta \leq 0,75)$.
- Två personer ska träffas på en viss plats mellan kl 12.00 och 13.00 och väntar högst 20 minuter på varandra. Om de anländer oberoende av varandra vid slumpmässigt valda tider under timmen, hur stor är sannolikheten att de träffas?

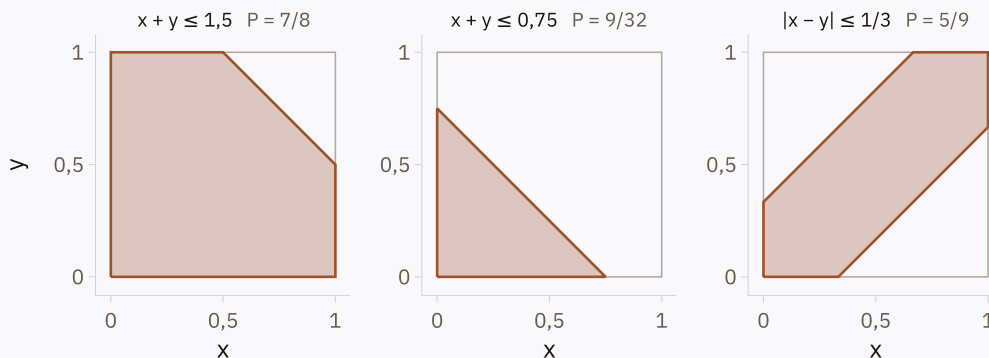
Lösning

- a)** Oberoende ger produkten av marginalerna, $f_{\xi}(x) = 1$ på $(0, 1)$ och $f_{\eta}(y) = 1$ på $(0, 1)$:

$$f(x, y) = f_{\xi}(x) f_{\eta}(y) = \begin{cases} 1, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{annars} \end{cases}$$

Det är den **likformiga fördelningen på enhetskvadraten** K , med $|K| = 1$.

- b)** $C = \{(x, y) : x + y \leq 1,5\}$, halvplanet under linjen $y = 1,5 - x$. Inom kvadraten är det allt utom ett hörn uppe till höger.



c) Komplementet inom kvadraten är triangeln med hörn i $(0,5; 1)$, $(1; 1)$ och $(1; 0,5)$, med kateterna 0,5:

$$P(\xi + \eta \leq 1,5) = 1 - \frac{0,5 \cdot 0,5}{2} = 1 - \frac{1}{8} = \boxed{\frac{7}{8}}$$

d) Nu är området själv en triangel med kateterna 0,75 (mittenpanelen):

$$P(\xi + \eta \leq 0,75) = \frac{0,75^2}{2} = \boxed{\frac{9}{32}}$$

e) Mät tiden i timmar efter kl 12: ankomsttiderna ξ och η är oberoende och $R(0, 1)$, alltså likformigt fördelade på kvadraten. De träffas om $|\xi - \eta| \leq 1/3$, bandet längs diagonalen i högra panelen. Utanför bandet ligger två trianglar med kateterna $2/3$:

$$P(|\xi - \eta| \leq \frac{1}{3}) = 1 - 2 \cdot \frac{(2/3)^2}{2} = 1 - \frac{4}{9} = \boxed{\frac{5}{9}}$$

Facit: a) $f(x, y) = 1$ om $0 \leq x, y \leq 1$, 0 annars; b) $C = \{(x, y) : x + y \leq 1,5\}$; c) $7/8$; d) $9/32$; e) $5/9$.

≡ Kompendiet, övning 1.1 – Seminarier och tentamensbetyg

För en universitetskurs fördes statistik över tentamensresultat och närvaro vid två frivilliga seminarier. Låt ξ vara betyget på tentamen (2 betyder underkänt) och η antalet seminarier som en slumpmässigt vald student har gått på. Den simultana sannolikhetsfunktionen $p(x, y)$ är

$x \backslash y$	0	1	2
2	0,18	0,07	0,05
3	0,11	0,16	0,08
4	0,06	0,08	0,11
5	0,02	0,03	0,05

Beräkna sannolikheten att en slumpmässigt vald student går på minst ett seminarium och får betyg 3 eller bättre.

Lösning

Området. Händelsen är $\{\xi \geq 3, \eta \geq 1\}$: raderna $x = 3, 4, 5$ och kolumnerna $y = 1, 2$, alltså sex celler i tabellens nedre högra hörn.

Summera massan där.

$$P(\xi \geq 3, \eta \geq 1) = (0,16 + 0,08) + (0,08 + 0,11) + (0,03 + 0,05) = \boxed{0,51}$$

Rimlighetskontroll. $P(\xi \geq 3) = 0,70$ och $P(\eta \geq 1) = 0,63$. Vore variablerna oberoende skulle svaret bli $0,70 \cdot 0,63 \approx 0,44$. Vi får mer, eftersom seminarienärvaro och bra betyg hänger ihop.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

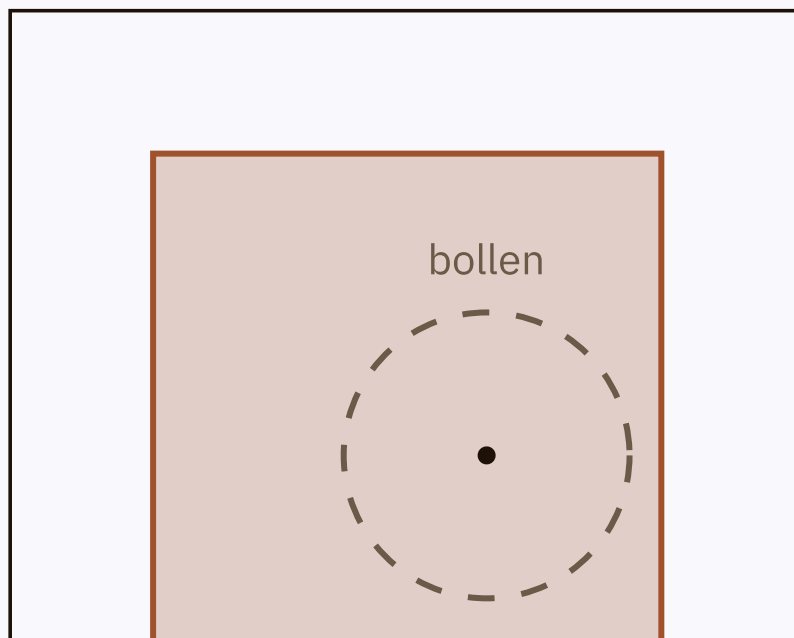
Kompendiet, övning 1.3 – Tennisboll genom ett nät

En tennisboll med diametern a cm kastas mot ett volleybollnät. Nätmaskorna är kvadratiska med sidan $b > a$. Vad är sannolikheten att bollen går igenom nätet utan att röra någon tråd? (Vi bortser från trådarnas tjocklek.)

Lösning

Modell. Bollens mittpunkt (ξ, η) hamnar i någon maska. Inom den maskan antar vi att mittpunkten är likformigt fördelad på kvadraten K med sidan b , så $|K| = b^2$.

Händelsen. Bollen rör ingen tråd om mittpunkten ligger minst en radie, $a/2$, från alla fyra sidor. Det är en inre kvadrat med sidan $b - a$:



Areaförhållandet.

$$P(\text{igenom}) = \frac{(b-a)^2}{b^2} = \left(1 - \frac{a}{b}\right)^2$$

Rimlighetskontroll. $a \rightarrow 0$ ger sannolikheten 1 och $a \rightarrow b$ ger 0, som sig bör.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

≡ Kompendiet, övning 1.6 – Vilken butik får första kunden?

En butik i Stockholm öppnar kl 9.00, och tiden tills den första kunden kommer in är exponentialfördelad med väntevärdet 4 minuter. En butik i Umeå öppnar också kl 9.00, med exponentialfördelad tid till första kunden med väntevärdet 7 minuter. Vad är sannolikheten att butiken i Stockholm får sin första kund före butiken i Umeå?

Lösning

Modell. ξ = tiden till första kunden i Stockholm, $\xi \in \text{Exp}(1/4)$; η = tiden i Umeå, $\eta \in \text{Exp}(1/7)$. Butikerna påverkar inte varandra, så ξ och η är oberoende och $f(x, y) = \frac{1}{4}e^{-x/4} \cdot \frac{1}{7}e^{-y/7}$ för $x, y > 0$.

Händelsen är området $\{(x, y) : x < y\}$ ovanför diagonalen. Integrera först över y från x till ∞ , vilket ger $P(\eta > x) = e^{-x/7}$:

$$P(\xi < \eta) = \int_0^{\infty} \frac{1}{4} e^{-x/4} e^{-x/7} dx = \frac{1}{4} \cdot \frac{1}{\frac{1}{4} + \frac{1}{7}} = \boxed{\frac{7}{11}} \approx 0,636$$

Tolkning. För två oberoende exponentialfördelningar gäller allmänt $P(\xi < \eta) = \lambda_{\xi}/(\lambda_{\xi} + \lambda_{\eta})$. Stockholm har den högre intensiteten (1/4 kund per minut mot 1/7) och vinner därför oftare.

På räknaren. $\int((1/4)*e^{(-x/4)}*e^{(-x/7)}, x, 0, \infty)$ ger $7/11$ i en Calculator-sida.

Ej kontrollerat mot facit – två oberoende lösningar stämmer överens.

≡ Kompendiet, övning 1.8 – Var ska sjukhuset ligga?

Ett sjukhus ska placeras i ett kvadratisk samhälle $K = \{(x, y) : 0 \leq x \leq 1, 0 \leq y \leq 1\}$, där $(0, 0)$ är sydvästra hörnet. Man kan bara färdas parallellt med koordinataxlarna, så avståndet från $(0, 0)$ till (x, y) är $x + y$. Antalet olyckor ökar linjärt i både x - och y -led, och platsen för en olycka har frekvensfunktionen

$$f(x, y) = \begin{cases} cxy, & 0 \leq x \leq 1, 0 \leq y \leq 1 \\ 0, & \text{annars} \end{cases}$$

- a) Bestäm konstanten c .
- b) Sjukhuset placeras i $(0, 0)$. Vad är sannolikheten att ambulansen behöver färdas högst 0,75?
- c) Vad är den förväntade utryckningssträckan med sjukhuset i $(0, 0)$?
- d) ✱ Vilken placering minimerar den förväntade utryckningssträckan?

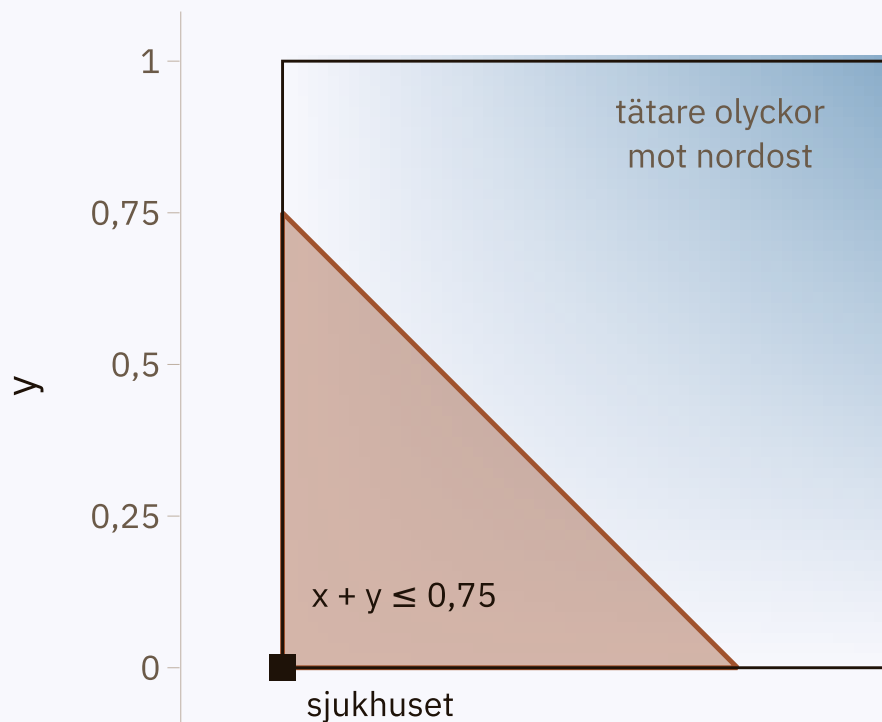
 Lösning

- a) Totala sannolikheten ska vara 1:

$$\int_0^1 \int_0^1 cxy \, dy \, dx = c \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{c}{4} = 1 \quad \Rightarrow \quad \boxed{c = 4}$$

Frekvensfunktionen faktoriserar, $4xy = 2x \cdot 2y$, så olycksplatsens koordinater ξ och η är **oberoende**, var och en med frekvensfunktionen $2x$ på $[0, 1]$.

- b) Händelsen är triangeln $\{x + y \leq 0,75\}$ i sydvästra hörnet, där olyckorna är som glesast:



$$P(\xi + \eta \leq 0,75) = \int_0^{0,75} \int_0^{0,75-x} 4xy \, dy \, dx = \int_0^{0,75} 2x(0,75-x)^2 \, dx = \frac{0,75^4}{6} = \boxed{\frac{27}{512}} \approx 0$$

Triangeln täcker 28 % av kvadratens area men bara 5 % av sannolikheten.

c) Väntevärdets linjäritet och $E(\xi) = \int_0^1 x \cdot 2x \, dx = \frac{2}{3}$:

$$E(\xi + \eta) = E(\xi) + E(\eta) = \boxed{\frac{4}{3}}$$

d) Med sjukhuset i (u, v) är sträckan $|\xi - u| + |\eta - v|$, och de två termerna kan minimeras var för sig. $E|\xi - u|$ minimeras när u är **medianen** för ξ – då ligger lika mycket sannolikhet på varje sida, och en liten flytt åt något håll förlänger lika mycket som den förkortar. $F_\xi(u) = u^2 = 1/2$ ger

$$(u, v) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}} \right) \approx (0,71; 0,71)$$

På räknaren. b) $\int (\int (4xy, y, 0, 0.75-x), x, 0, 0.75)$ ger 27/512.

Ej kontrollerat mot facit — två oberoende lösningar stämmer överens.

Se även

- Marginalfördelning
- Betingad fördelning och betingat väntevärde
- Kovarians
- Korrelation
- Tvådimensionell normalfördelning
- Dubbelintegraler

Resurser

- [Khan Academy: Joint distributions](#) 
- [Wikipedia: Joint probability distribution](#) 