Sannolikhet och statistik

Flerdimensionella stokastiska variabler

VT 2009

Uwe.Menzel@math.uu.se

http://www.math.uu.se/∼uwe/

Flerdimensioneller slumpvariabler

Ett slumpförsök kan ge upphov till två (eller fler) reella tal.

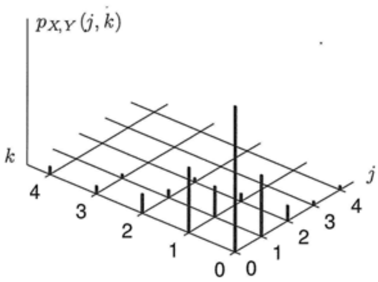
Två tal "emitteras" vid varje utfall: $\Omega \Rightarrow \mathbb{R}^2$

Exempel

- kast med två tärningar
- slumpmässigt vald person: (vikt, längd)
- Buffons nålproblem: vinkel och avstånd till planskarv $\Rightarrow \pi$

Sannolikhetsfunktion för en diskret 2-dim. slumpvariabel

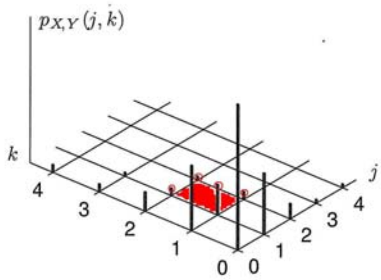
$$\boxed{p_{X,Y}(j,k) = P(X = j, Y = k)}$$



Figur: Barnkullar, Blom s. 86

Diskret 2-dim. slumpvariabel: Sannolikhet för händelser

$$P(X, Y \in A) = \sum_{(j,k \in A)} P_{X,Y}(j,k)$$



Figur: Familier med 1 eller 2 flickor och 1 eller 2 pojkar

Sannolikhetsfunktion: Egenskaper

$$p_{X,Y}(k,j) \geq 0$$

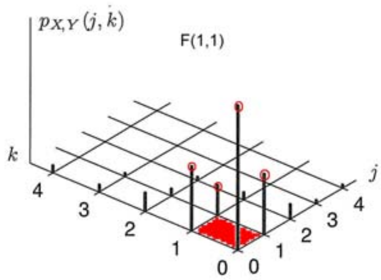
$$\sum_k \sum_j p_{X,Y}(k,j) = 1$$

$j \backslash k$	0	1	2	3	4	Summa
0	0.38	0.16	0.04	0.01	0.01	0.60
1	0.17	0.08	0.02			0.27
2	0.05	0.02	0.01			0.08
3	0.02	0.01				0.03
4	0.02					0.02
Summa	0.64	0.27	0.07	0.01	0.01	$\boxed{1.00}$

Figur: Barnkullar, Normalisering

Diskret 2-dim. slumpvariabel: Fördelningsfunktion

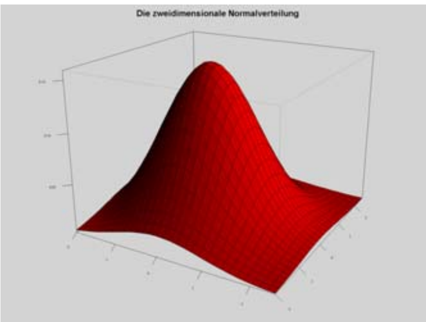
$$P(X \leq x, Y \leq y) = F_{X,Y}(x,y) = \sum_{j \leq x} \sum_{k \leq y} p_{X,Y}(j,k)$$



Figur: Fördelningsfunktion: summa över rektangeln

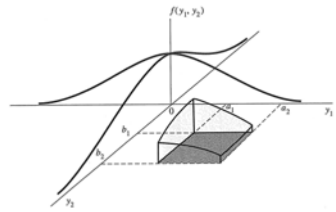
Täthetsfunktion för kontinuerlig 2-dim. slumpvariabel

$$\boxed{f_{X,Y}(x,y)}$$



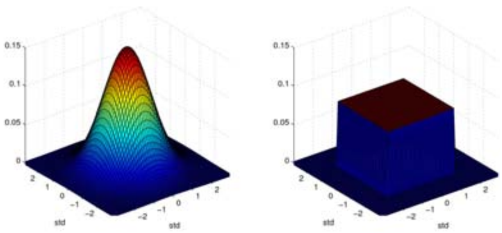
Sannolikhet för händelser

$$P(a_1 < X \leq a_2, b_1 < Y \leq b_2) = \int_{b_1}^{b_2} \int_{a_1}^{a_2} f_{X,Y}(x,y)\, dx\, dy$$



Tvådimensionell likformig fördelning

$$P(X \in A) = \frac{arean\ av\ A}{arean\ av\ \Omega}$$

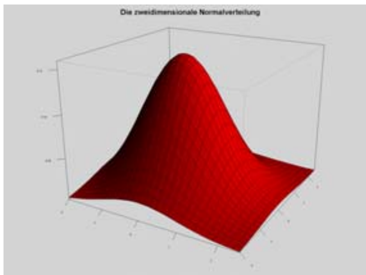


Figur: Tvådimensionell normalfördelning och likformig fördelning

Täthetsfunktion: Egenskaper

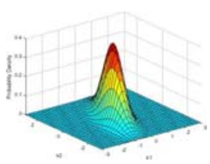
$$f_{X,Y}(x,y) \geq 0$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f_{X,Y}(x,y)\, dx\, dy = 1$$

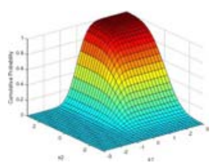


Fördelningsfunktion

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



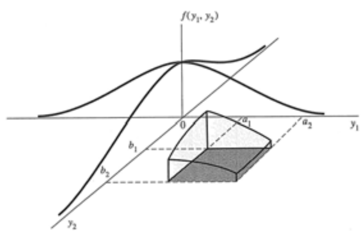
Figur: Täthetsfunktion



Figur: Fördelningsfunktion

Sannolikhet för händelser med hjälp av fördelningsfunktionen

$$\begin{aligned} P(a_1 < X \leq a_2, b_1 < Y \leq b_2) &= F_{X,Y}(a_2, b_2) \\ &- F_{X,Y}(a_2, b_1) \\ &- F_{X,Y}(a_1, b_2) \\ &+ F_{X,Y}(a_1, b_1) \end{aligned}$$



Diskret 2-dim. slumpvariabel: Marginalfördelning

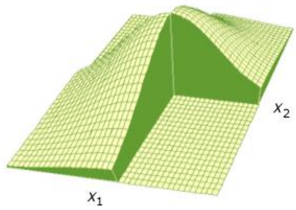
$$p_X(k) = \sum_j^{\infty} p_{X,Y}(k,j)$$
$$p_Y(j) = \sum_k^{\infty} p_{X,Y}(k,j)$$

$j \backslash k$	0	1	2	3	4	Summa
0	0.38	0.16	0.04	0.01	0.01	0.60
1	0.17	0.08	0.02			0.27
2	0.05	0.02	0.01			0.08
3	0.02	0.01				0.03
4	0.02					0.02
Summa	0.64	0.27	0.07	0.01	0.01	1.00

Figur: Barnkullar, Marginalfördelningar

Kontinuerlig 2-dim. slumpvariabel: Marginalfördelning

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



Oberoende stokastiska variabler

Slumpvariablerna X och Y är oberoende om:

$$F_{X,Y}(x,y) = F_X(x) \cdot F_Y(y) \quad \text{för alla x och y}$$

eller

$$p_{X,Y}(j,k) = p_X(j) \cdot p_Y(k) \quad \text{för alla j och k, diskret s.v.}$$
$$f_{X,Y}(x,y) = f_X(x) \cdot f_Y(y) \quad \text{för alla x och y, kontin. s.v.}$$

Kom ihåg: $P(A \cap B) = P(A) \cdot P(B)$ för oberoende händelser

Oberoende stokastiska variabler

Tärningskast:

$$P(X=1, Y=6) = P(X=1) \cdot P(Y=6) = \frac{1}{6} \cdot \frac{1}{6} = \frac{1}{36}$$

Barnkullar:

$j \backslash k$	0	1	2	3	4	Summa
0	0.38	0.16	0.04	0.01	0.01	0.60
1	0.17	0.08	0.02			0.27
2	0.05	0.02	0.01			0.08
3	0.02	0.01				0.03
4	0.02					0.02
Summa	0.64	0.27	0.07	0.01	0.01	1.00

Figur: $p_{X,Y}(j,k) = p_X(k) \cdot p_Y(j)$??