χ^2 raspodjela

Def:

Kontinuirana slučajna varijabla X ima χ^2 **raspodjelu** s parametrom $n \in \mathbb{N}$ ako je njezina funkcija gustoće vjerojatnosti dana s

$$f(x;n)=\begin{cases}\dfrac{x^{n/2-1}e^{-x/2}}{2^{n/2}\Gamma(n/2)} & , \quad x\geq 0\\[2ex] 0 & , \quad x<0\end{cases}$$

Parametar n zove se "broj stupnjeva slobode" varijable X.

$\alpha=n/2$; $\beta=2$

χ^2 raspodjela je važna za primjenu statističkih testova.

Dvodimenzionalne raspodjele

Def:

Neka su X i Y dvije diskretne slučajne varijable definirane na prostoru elementarnih događaja nekog eksperimenta. **Združena raspodjela vjerojatnosti** $p(x,y)$ definirana je za svaki par točaka (x,y) kao vjerojatnost da istodobno varijabla X poprimi vrijednost x i varijabla Y poprimi vrijednost y:

$$p(x,y)=P(X=x \text{ i } Y=y)$$

	M	Ž	
Matematika	0,151	0,105	
Fizika	0,125	0,079	
Kemija	0,069	0,074	
Biologija	0,062	0,153	
Geo-znanosti	0,124	0,058	

Def:

Rubne (marginalne) raspodjele vjerojatnosti za varijable X i Y označavamo $p_X(x)$ i $p_Y(y)$, a dane su izrazima:

$$p_X(x)=\sum_{y\in D_Y} p(x,y) \qquad i \qquad p_Y(y)=\sum_{x\in D_X} p(x,y)$$

	M	Ž	ukupno
Matematika	0,151	0,105	0,256
Fizika	0,125	0,079	0,204
Kemija	0,069	0,074	0,143
Biologija	0,062	0,153	0,215
Geo-znanosti	0,124	0,058	0,182
ukupno	0,531	0,469	1

Def:

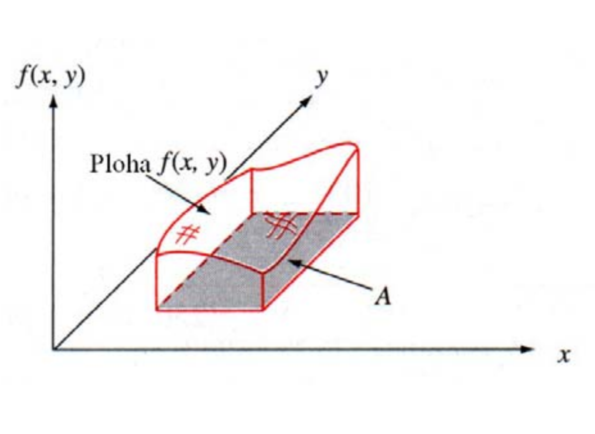
Neka su X i Y dvije kontinuirane slučajne varijable. Funkcija $f(x,y)$ je njihova **združena funkcija gustoće vjerojatnosti** ako za bilo koji dvodimenzionalni skup $A \subseteq \mathbb{R}\times\mathbb{R}$ vrijedi

$$P((X,Y)\in A)=\iint_A f(x,y)\,dxdy$$

Posebno, ako je A dvodimenzionalni pravokutnik $A=\{(x,y): a\leq x\leq b, c\leq y\leq d\}$, tada je

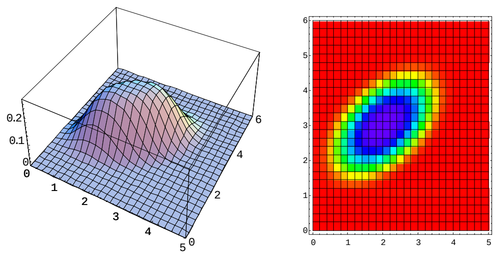
$$P((X,Y)\in A)=P(a\leq x\leq b, c\leq y\leq d)=\int_a^b\int_c^d f(x,y)\,dxdy$$

$$f(x,y)\geq 0 \qquad \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x,y)\,dxdy=1$$



PRIMJER

$$f(x,y)=\frac{\sqrt{3}}{4\pi}e^{-\frac{(x-2)^2+(y-3)^2-(x-2)(y-3)}{2}}$$



Def:

Rubne (marginalne) funkcije gustoće vjerojatnosti za varijable X i Y označavamo $f_X(x)$ i $f_Y(y)$, a dane su izrazima:

$$f_X(x)=\int_{-\infty}^{\infty} f(x,y)dy \qquad i \qquad f_Y(y)=\int_{-\infty}^{\infty} f(x,y)dx$$

Nezavisne slučajne varijable

Def:

Slučajne varijable X i Y su **nezavisne** ako za svaki par vrijednosti x i y vrijedi

$p(x,y)=p_X(x)\cdot p_Y(y)$ (za diskretne varijable)

ili

$f(x,y)=f_X(x)\cdot f_Y(y)$ (za kontinuirane varijable).

Ako ti uvjeti nisu ispunjeni za sve (x,y), kažemo da su X i Y **zavisne**.

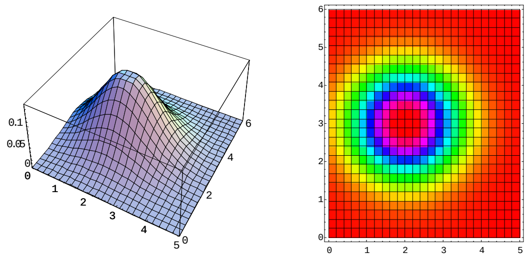
$$P(A\cap B)=P(A)\cdot P(B)$$

Provjeriti nezavisnost!

	M	Ž	ukupno
Matematika	0,151	0,105	0,256
Fizika	0,125	0,079	0,204
Kemija	0,069	0,074	0,143
Biologija	0,062	0,153	0,215
Geo-znanosti	0,124	0,058	0,182
ukupno	0,531	0,469	1

Calculator

$$f(x,y)=\frac{1}{2\pi}e^{-\frac{(x-2)^2+(y-3)^2}{2}}$$



Očekivanja

$$\mu_X = \sum_{x\in D_X} x\cdot p_X(x) = \sum_{x\in D_X}\sum_{y\in D_Y} x\cdot p(x,y) \qquad ili$$

$$\mu_X = \int_{-\infty}^{\infty} x\cdot f_X(x)dx = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\cdot f(x,y)dxdy$$

$$\mu_Y = \sum_{y\in D_Y} y\cdot p_Y(y) = \sum_{x\in D_X}\sum_{y\in D_Y} y\cdot p(x,y) \qquad ili$$

$$\mu_Y = \int_{-\infty}^{\infty} y\cdot f_Y(y)dy = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} y\cdot f(x,y)dxdy$$

Momenti 2D raspodjele

Pomoćni momenti

$$m_{rs} = \sum_{x\in D_X}\sum_{y\in D_Y} x^r y^s p(x,y) \qquad ili$$

$$m_{rs} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x^r y^s f(x,y)dxdy$$

Središnji momenti

$$M_{rs} = \sum_{x\in D_X}\sum_{y\in D_Y} (x-\mu_X)^r (y-\mu_Y)^s p(x,y) \qquad ili$$

$$M_{rs} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x-\mu_X)^r (y-\mu_Y)^s f(x,y)dxdy$$

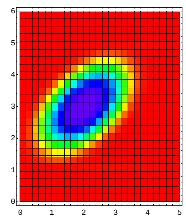
$m_{10} = \mu_X$

$m_{01} = \mu_Y$ očekivanja

$M_{20} = V(X) = \sigma_X^2$

$M_{02} = V(Y) = \sigma_Y^2$ varijance

$M_{11} = \sigma_{XY} = Cov(X,Y)$ kovarijanca



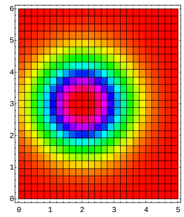
$m_{10} = 2 = \mu_X$

$m_{01} = 3 = \mu_Y$ očekivanja

$M_{20} = V(X) = \frac{4}{3}$

$M_{02} = V(Y) = \frac{4}{3}$ varijance

$M_{11} = \sigma_{XY} = \frac{2}{3}$ kovarijanca



$m_{10} = 2 = \mu_X$

$m_{01} = 3 = \mu_Y$ očekivanja

$M_{20} = V(X) = 1$

$M_{02} = V(Y) = 1$ varijance

$M_{11} = \sigma_{XY} = 0$ kovarijanca

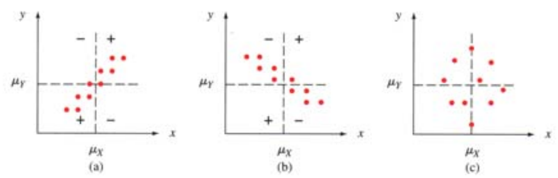


Figure 5.4 $p(x, y) = 1/10$ for each of ten pairs corresponding to indicated points; (a) positive covariance; (b) negative covariance; (c) covariance near zero

Pomoćna formula

$$Cov(X,Y) = m_{11} - m_{10}\cdot m_{01} = E(X\cdot Y) - E(X)E(Y)$$

Korelacija

koeficijent korelacije: $Corr(X,Y) = \rho_{X,Y} = \rho = \dfrac{\sigma_{XY}}{\sigma_X\cdot\sigma_Y}$

Svojstva:

1. $Corr(aX+b, cY+d) = Corr(X,Y)$
2. $-1 \le \rho \le 1$
3. Ako su X i Y nezavisne varijable, onda je ρ=0. (Obrat ne vrijedi)
4. Ako je Y=aX+b onda i samo onda je ρ=1 ili ρ= -1.

Linearna kombinacija dviju slučajnih varijabli

2D raspodjela

Slučajne varijable X_1 i X_2 koje se mogu zbrajati

Združena raspodjela vjerojatnosti $p(X_1,X_2)$

$$a_1, a_2 \in \boldsymbol{R}$$

Def: $Y = a_1X_1 + a_2X_2$ (linearna kombinacija)

Očekivanje linearne kombinacije

$$E(Y) = a_1E(X_1) + a_2E(X_2)$$

Varijanca linearne kombinacije

$$V(Y) = a_1^2V(X_1) + a_2^2V(X_2) + 2a_1a_2Cov(X_1, X_2)$$

Varijanca razlike nezavisnih varijabli

$$V(X_1 - X_2) = V(X_1) + V(X_2) = \sigma_1^2 + \sigma_2^2$$

Varijanca razlike je zbroj varijanci, a ne razlika!