5. Simultana fördelningar

Antag att man har två diskreta stokastiska variabler X och Y. Den **simultana frekvensfunktionen** av X och Y definieras

$$f(x,y) = P(X = x, Y = y)$$

Funktionen är sådan att

$$\sum_{\text{alla } y} \sum_{\text{alla } x} P(X = x, Y = y) = 1$$

Frekvensfunktionerna av X och Y får man från den simultana frekvensfunktionen:

$$f_X(x) = P(X = x) = \sum_{\text{alla } y} P(X = x, Y = y)$$

och

$$f_Y(y) = P(Y = y) = \sum_{\text{alla } x} P(X = x, Y = y)$$

Om X och Y är oberoende, då är

$$P(X = x, Y = y) = P(X = x)P(Y = y)$$

för alla x och y.

Obs! I kontinuerliga fallet kan man definiera den **simultana täthetsfunktionen**.

Man kan också räkna den betingade sannolikheten för en viss x givet att Y får ett värde y:

$$P(X = x|Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

om $P(Y = y) > 0$.

Låt X vara en stokastisk variabel med väntevärdet μ_X och Y en stokastisk variabel med väntevärdet μ_Y. **Kovariansen** mellan X och Y är

$$Cov(X,Y) = \mathbf{E}[(X - \mu_X)(Y - \mu_Y)] = \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = \mathbf{E}[XY] - \mu_X \mu_Y.$$

Väntevärdet $\mathbf{E}[XY]$ kan man räkna (i diskreta fallet) genom att använda den simultana frekvensfunktionen

$$\mathbf{E}[XY] = \sum_{\text{alla } y} \sum_{\text{alla } x} xy\, P(X = x, Y = y)$$

Korrelationen mellan X och Y mäter linjärt beroende mellan X och Y. Korrelationen mellan X och Y är

$$\rho_{XY} = \frac{Cov(X,Y)}{\sigma_X \sigma_Y}$$

där σ_X och σ_Y är respektive standardavvikelser av X och Y.