STAT101 Elementær statistikk høsten 2021

- Addisjonsregelen: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- Betinget sannsynlighet for A gitt B: $P(A|B) = \frac{P(A \cap B)}{P(B)}$
- Bayes teorem: $P(B|A) = \frac{P(A|B)P(B)}{P(A)}$
- Loven om total sannsynlighet: Dersom $B_1, \ldots, B_n$ er disjunkte hendelser og $\cup_{i=1}^n B_n = S$, hvor S er utfallsrommet

$$P(A) = P(A \cap B_1) + \cdots + P(A \cap B_n) = P(A|B_1)P(B_1) + \cdots + P(A|B_n)P(B_n)$$

- For en diskret tilfeldig variabel X er forventningsverdien til X definert ved:

$$\mu_X = \sum_{x \in V_X} xP(X = x)$$

- For en tilfeldig variabel X er variansen til X definert ved:

$$\sigma_X^2 = \mu_{(X-\mu_X)^2}$$

- For vilkårlige konstanter a og b og to tilfeldige variabler X og Y gjelder:

$$\sigma_{(aX+bY)}^2 = a^2\sigma_X^2 + b^2\sigma_Y^2 + 2ab \cdot \rho\sigma_X\sigma_Y$$

  hvor ρ er korrelasjonen mellom X og Y.

- For vilkårlige konstanter $a_1, a_2, \ldots, a_n$ og b og tilfeldige variable $X_1, \ldots, X_n$ gjelder:

$$\mu_{a_1X_1+a_2X_2+\cdots+a_nX_n+b} = a_1\mu_{X_1} + a_2\mu_{X_2} + \cdots + a_n\mu_{X_n} + b$$

  Dersom $X_1, \ldots, X_n$ i tillegg er uavhengige tilfeldige variabler gjelder også:

$$\sigma^2_{a_1X_1+a_2X_2+\cdots+a_nX_n+b} = a_1^2\sigma_{X_1}^2 + a_2^2\sigma_{X_2}^2 + \cdots + a_n^2\sigma_{X_n}^2$$

- Dersom den tilfeldige variabelen X er binomisk fordelt med parametre n og p har vi at:

$$P(X = x) = \binom{n}{x} p^x(1-p)^{n-x}, \qquad x = 0, 1, \ldots, n,$$

$$\mu_X = np \text{ og } \sigma_X^2 = np(1-p).$$

- Sentralgrenseteoremet:

  Anta $X_1, \ldots, X_n$ er uavhengige og identisk fordelte tilfeldige variabler med forventning μ og varians σ^2. For z et tall på tallinjen og n "tilstrekkelig stor":

$$P\left(\frac{S_n - n\mu}{\sqrt{n}\sigma} \leq z\right) \approx P(Z \leq z) \quad \text{og} \quad P\left(\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \leq z\right) \approx P(Z \leq z), \quad \text{der } Z \sim N(0,1),$$

  hvor $S_n = X_1 + \cdots + X_n$ og $\bar{X} = \frac{1}{n}\sum_{i=1}^n X_i$.

- For vilkårlige konstanter $a_1, a_2, \ldots, a_n$, dersom $X_1, \ldots, X_n$ er uavhengige normalfordelte tilfeldige variable, gjelder det at også $a_1X_1 + a_2X_2 + \cdots + a_nX_n$ er normalfordelt.

- Ett-utvalgs testobservator:

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t(n-1), \quad \text{hvor } S = \frac{1}{(n-1)} \sum_{i=1}^{n} (X_i - \bar{X})^2$$

- Pooled to-utvalgs testobservator:

$$T = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{S_p \sqrt{\frac{1}{n_X} + \frac{1}{n_Y}}} \sim t_{n_X + n_Y - 2}, \quad \text{hvor} \quad S_p^2 = \frac{(n_X - 1)S_X^2 + (n_Y - 1)S_Y^2}{n_X + n_Y - 2}.$$

- To-utvalgs testobservator:

$$T = \frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{\sqrt{\frac{S_X^2}{n_X} + \frac{S_Y^2}{n_Y}}} \sim t_m, \quad \text{hvor} \quad m = \min(n_X - 1, n_Y - 1)$$

- Lineær regresjon:
  - $\hat{y} = a + bx$, der $b = r\frac{s_y}{s_x}$ og $a = \bar{y} - b\bar{x}$
  - $s^2 = \frac{1}{n-2} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2$.
  - $SE_b = \frac{s}{\sum_{i=1}^{n} (x_i - \bar{x})^2}$
  - Et $(1 - \alpha)100\%$ konfidensintervall for μ_y ved gitt x-verdi lik x^*:

$$a + bx^* \pm t_{\alpha/2, n-2} \cdot s \cdot \sqrt{\frac{1}{n} + \frac{(x^* - \bar{x})^2}{(s/SE_b)^2}}$$

  - Et $(1 - \alpha)100\%$ prediksjonsintervall for y-verdien ved gitt x-verdi lik x^*:

$$a + bx^* \pm t_{\alpha/2, n-2} \cdot s \cdot \sqrt{1 + \frac{1}{n} + \frac{(x^* - \bar{x})^2}{(s/SE_b)^2}}$$