

FORMELSAMLING TIL STK1100 OG STK1110

(Versjon av 11. november 2017)

1. Sannsynlighet

La $A, B, A_1, A_2, \dots, B_1, B_2, \dots$ være begivenheter, dvs. delmengder av et utfallsrom Ω .

a) Aksiomene:

Et sannsynlighetsmål P er en funksjon fra delmengder av utfallsrommet Ω til de reelle tall som tilfredsstiller

$$P(\Omega) = 1$$

$$P(A) \geq 0$$

$$P(A_1 \cup A_2) = P(A_1) + P(A_2) \quad \text{hvis } A_1 \cap A_2 = \emptyset$$

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} P(A_i) \quad \text{hvis } A_i \cap A_j = \emptyset \text{ for } i \neq j$$

b) $P(A') = 1 - P(A)$

c) $P(\emptyset) = 0$

d) $A \subset B \Rightarrow P(A) \leq P(B)$

e) Addisjonsetningen:

$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$

f) Betinget sannsynlighet:

$$P(A|B) = \frac{P(A \cap B)}{P(B)} \quad \text{hvis } P(B) > 0$$

g) Total sannsynlighet:

$$P(A) = \sum_{i=1}^n P(A|B_i)P(B_i) \quad \text{hvis } \bigcup_{i=1}^n B_i = \Omega \text{ og } B_i \cap B_j = \emptyset \text{ for } i \neq j$$

h) Bayes' setning:

$$P(B_j|A) = \frac{P(A|B_j)P(B_j)}{\sum_{i=1}^n P(A|B_i)P(B_i)} \quad \text{under samme betingelser som i g)}$$

i) A og B er uavhengige begivenheter hvis $P(A \cap B) = P(A)P(B)$

j) A_1, \dots, A_n er uavhengige begivenheter dersom

$$P(A_{i_1} \cap \dots \cap A_{i_m}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_m})$$

for alle delmengder av indekser i_1, i_2, \dots, i_m

k) Produktsetningen:

$$\begin{aligned} P(A_1 \cap \dots \cap A_n) \\ = P(A_1)P(A_2|A_1)P(A_3|A_1 \cap A_2) \cdots P(A_n|A_1 \cap A_2 \cap \dots \cap A_{n-1}) \end{aligned}$$

2. Kombinatorikk

- a) To operasjoner som kan gjøres på henholdsvis n og m måter kan kombineres på $n \cdot m$ måter.
- b) Antall ordnete utvalg med tilbakelegging av r elementer fra en mengde med n elementer er n^r
- c) Antall ordnete utvalg uten tilbakelegging av r elementer fra en mengde med n elementer er $n(n-1) \cdots (n-r+1)$
- d) Antall måter n elementer kan ordnes i rekkefølge på (permuteres) er $n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$
- e) Antall ikke-ordete utvalg av r elementer fra en mengde med n elementer er

$$\binom{n}{r} = \frac{n(n-1) \cdots (n-r+1)}{r!} = \frac{n!}{r!(n-r)!}$$

- f) Antall måter en mengde med n elementer kan deles inn i r delmengder med n_i elementer i den i -te delmengden er

$$\binom{n}{n_1 \ n_2 \ \cdots \ n_r} = \frac{n!}{n_1! n_2! \cdots n_r!}$$

3. Sannsynlighetsfordelinger

- a) For en stokastisk variabel X (diskret eller kontinuert) er den kumulative fordelingsfunksjonen $F(x) = P(X \leq x)$
- b) For en diskret stokastisk variabel X som kan anta verdiene x_1, x_2, x_3, \dots har vi

$$p(x_j) = P(X = x_j)$$
$$F(x) = \sum_{x_j \leq x} p(x_j)$$

Betingelsene for at $p(x_j)$ skal være en punktsannsynlighet er

$$p(x_j) \geq 0 \quad \text{for alle } j$$
$$\sum_j p(x_j) = 1$$

- c) For en kontinuert stokastisk variabel X har vi

$$P(a < X < b) = \int_a^b f(x) dx$$
$$F(x) = \int_{-\infty}^x f(u) du$$
$$f(x) = F'(x)$$

Betingelsene for at $f(x)$ skal være en sannsynlighetstetthet er

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

- d) For to stokastiske variabler X og Y (diskrete eller kontinuerte) er den simultane kumulative fordelingsfunksjonen $F(x, y) = P(X \leq x, Y \leq y)$
- e) For diskrete stokastiske variabler X og Y som kan anta henholdsvis verdiene x_1, x_2, \dots og y_1, y_2, \dots har vi

$$p(x_i, y_j) = P(X = x_i, Y = y_j)$$
$$F(x, y) = \sum_{x_i \leq x} \sum_{y_j \leq y} p(x_i, y_j)$$

Betingelsene for at $p(x_i, y_j)$ skal være en simultan punktsannsynlighet er analoge til betingelsene i b)

f) For kontinuerlige stokastiske variabler X og Y har vi

$$P((X, Y) \in A) = \int \int_A f(u, v) dv du$$

$$F(x, y) = \int_{-\infty}^x \int_{-\infty}^y f(u, v) dv du$$

$$f(x, y) = \frac{\partial^2 F(x, y)}{\partial x \partial y}$$

Betingelsene for at $f(x, y)$ skal være en simultan sannsynlighetstetthet er analoge til betingelsene i c)

g) Marginale punktsannsynligheter:

$$p_X(x_i) = \sum_j p(x_i, y_j) \quad (\text{for } X)$$

$$p_Y(y_j) = \sum_i p(x_i, y_j) \quad (\text{for } Y)$$

h) Marginale sannsynlighetstettheter:

$$f_X(x) = \int_{-\infty}^{\infty} f(x, y) dy \quad (\text{for } X)$$

$$f_Y(y) = \int_{-\infty}^{\infty} f(x, y) dx \quad (\text{for } Y)$$

i) Uavhengighet:

De stokastiske variablene X og Y er uavhengige dersom

$$p(x_i, y_j) = p_X(x_i)p_Y(y_j) \quad (\text{diskret})$$

$$f(x, y) = f_X(x)f_Y(y) \quad (\text{kontinuerlig})$$

j) Betingete punktsannsynligheter:

$$p_{X|Y}(x_i|y_j) = \frac{p(x_i, y_j)}{p_Y(y_j)} \quad (\text{for } X \text{ gitt } Y = y_j)$$

$$p_{Y|X}(y_j|x_i) = \frac{p(x_i, y_j)}{p_X(x_i)} \quad (\text{for } Y \text{ gitt } X = x_i)$$

Det forutsettes at $p_Y(y_j) > 0$ og $p_X(x_i) > 0$, henholdsvis. De betingete punktsannsynlighetene kan behandles som vanlige punktsannsynligheter.

k) Betingete sannsynlighetstettheter:

$$f_{X|Y}(x|y) = \frac{f(x, y)}{f_Y(y)} \quad (\text{for } X \text{ gitt } Y = y)$$

$$f_{Y|X}(y|x) = \frac{f(x, y)}{f_X(x)} \quad (\text{for } Y \text{ gitt } X = x)$$

Det forutsettes at $f_Y(y) > 0$ og $f_X(x) > 0$, henholdsvis. De betingete sannsynlighetstetthetene kan behandles som vanlige sannsynlighetstettheter.

4. Forventning

a) Forventningsverdien til en stokastisk variabel X er definert ved

$$E(X) = \sum_j x_j p(x_j) \quad (\text{diskret})$$

$$E(X) = \int_{-\infty}^{\infty} x f(x) dx \quad (\text{kontinuerlig})$$

b) For en reell funksjon $g(X)$ av en stokastisk variabel X er

$$E[g(X)] = \sum_j g(x_j) p(x_j) \quad (\text{diskret})$$

$$E[g(X)] = \int_{-\infty}^{\infty} g(x) f(x) dx \quad (\text{kontinuerlig})$$

c) $E(a + bX) = a + bE(X)$

d) For en reell funksjon $g(X, Y)$ av to stokastiske variabler X og Y er

$$E[g(X, Y)] = \sum_i \sum_j g(x_i, y_j) p(x_i, y_j) \quad (\text{diskret})$$

$$E[g(X, Y)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x, y) f(x, y) dy dx \quad (\text{kontinuerlig})$$

e) Hvis X og Y er uavhengige er $E[g(X)h(Y)] = E[g(X)] \cdot E[h(Y)]$

f) Hvis X og Y er uavhengige er $E(XY) = E(X) \cdot E(Y)$

g) $E\left(a + \sum_{i=1}^n b_i X_i\right) = a + \sum_{i=1}^n b_i E(X_i)$

h) Betinget forventning:

$$E(Y|X = x_i) = \sum_j y_j p_{Y|X}(y_j|x_i) \quad (\text{diskret})$$

$$E(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x) dy \quad (\text{kontinuerlig})$$

5. Varians og standardavvik

a) Variansen og standardavviket til en stokastisk variabel X er definert ved

$$V(X) = E[(X - \mu)^2]$$
$$\text{sd}(X) = \sqrt{V(X)}$$

b) $V(X) = E(X^2) - (E(X))^2$

c) $V(a + bX) = b^2 V(X)$

d) Hvis X_1, \dots, X_n er uavhengige har vi

$$V\left(a + \sum_{i=1}^n b_i X_i\right) = \sum_{i=1}^n b_i^2 V(X_i)$$

e)

$$V\left(a + \sum_{i=1}^n b_i X_i\right) = \sum_{i=1}^n b_i^2 V(X_i) + \sum_{i=1}^n \sum_{j \neq i} b_i b_j \text{Cov}(X_i, X_j)$$

f) Chebyshev's ulikhet:

La X være en stokastisk variabel med $\mu = E(X)$ og $\sigma^2 = V(X)$.

For alle $t > 0$ har vi

$$P(|X - \mu| > t) \leq \frac{\sigma^2}{t^2}$$

6. Kovarians og korrelasjon

a) La X og Y være stokastiske variabler med $\mu_X = E(X)$, $\sigma_X^2 = V(X)$, $\mu_Y = E(Y)$ og $\sigma_Y^2 = V(Y)$. Da er kovariansen og korrelasjonen til X og Y definert ved

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

$$\rho = \text{Corr}(X, Y) = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

- b) $\text{Cov}(X, X) = V(X)$
 c) $\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$
 d) X, Y uavhengige $\Rightarrow \text{Cov}(X, Y) = 0$
 e)

$$\text{Cov}\left(a + \sum_{i=1}^n b_i X_i, c + \sum_{j=1}^m d_j Y_j\right) = \sum_{i=1}^n \sum_{j=1}^m b_i d_j \text{Cov}(X_i, Y_j)$$

- f) $-1 \leq \text{Corr}(X, Y) \leq 1$ og $\text{Corr}(X, Y) = \pm 1$ hvis og bare hvis det finnes to tall a, b slik at $Y = a + bX$ (bortsett, eventuelt, på et område med sannsynlighet 0)

7. Momentgenererende funksjoner

- a) For en stokastisk variabel X (diskret eller kontinuerlig) er den momentgenererende funksjonen $M_X(t) = E(e^{tX})$
 b) Hvis den momentgenererende funksjonen $M_X(t)$ eksisterer for t i et åpent intervall som inneholder null, så bestemmer den entydig fordelingen til X
 c) Hvis den momentgenererende funksjonen $M_X(t)$ eksisterer for t i et åpent intervall som inneholder null, så eksisterer alle momenter til X , og vi kan finne det r -te momentet ved $E(X^r) = M_X^{(r)}(0)$
 d) $M_{a+bX}(t) = e^{at}M_X(bt)$
 e) Hvis X og Y er uavhengige er $M_{X+Y}(t) = M_X(t)M_Y(t)$

8. Noen diskrete sannsynlighetsfordelinger

- a) Binomisk fordeling:

Punktsannsynlighet: $P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad k = 0, 1, \dots, n$

Momentgenererende funksjon: $M_X(t) = (1-p + pe^t)^n$

Forventning: $E(X) = np$

Varians: $V(X) = np(1-p)$

Tilnærmelse 1: $Z = \frac{X - np}{\sqrt{np(1-p)}}$ er tilnærmet standard normalfordelt når np og $n(1-p)$ begge er tilstrekkelig store (minst 10)

Tilnærmelse 2: X er tilnærmet Poisson fordelt med parameter $\lambda = np$ når n er stor og p er liten

Addisjonsregel: $X \sim \text{binomisk}(n, p)$, $Y \sim \text{binomisk}(m, p)$
og X, Y uavhengige $\Rightarrow X + Y \sim \text{binomisk}(n + m, p)$

b) Geometrisk fordeling:

Punktsannsynlighet: $P(X = k) = (1 - p)^{k-1}p \quad k = 1, 2, \dots$

Momentgenererende funksjon : $M_X(t) = e^t p / [1 - (1 - p)e^t]$

Forventning: $E(X) = 1/p$

Varians: $V(x) = (1 - p)/p^2$

Addisjonsregel: Hvis X er geometrisk fordelt med sannsynlighet p så er $X - 1$ negativt binomisk $(1, p)$. Derfor hvis X og Y er geometrisk fordelte med samme p og uavhengige så er $X + Y - 2$ negativt binomisk $(2, p)$

c) Negativ binomisk fordeling:

Punktsannsynlighet: $P(X = k) = \binom{k+r-1}{r-1} p^r (1 - p)^k \quad k = 0, 1, 2, \dots$

Momentgenererende funksjon : $M_X(t) = \{p/[1 - (1 - p)e^t]\}^r$

Forventning: $E(X) = r(1 - p)/p$

Varians: $V(X) = r(1 - p)/p^2$

Addisjonsregel: $X \sim \text{negativ binomisk}(r_1, p)$,
 $Y \sim \text{negativt binomisk}(r_2, p)$
og X, Y uavhengige
 $\Rightarrow X + Y \sim \text{negativt binomisk}(r_1 + r_2, p)$

d) Hypergeometrisk fordeling:

Punktsannsynlighet: $P(X = k) = \frac{\binom{M}{k} \binom{N-M}{n-k}}{\binom{N}{n}}$

Forventning: $E(X) = n \cdot \frac{M}{N}$

Varians: $V(X) = n \frac{M}{N} (1 - \frac{M}{N}) \frac{N-n}{N-1}$

Tilnærmelse: X er tilnærmet binomisk $(n, \frac{M}{N})$ når n er mye mindre enn N

e) Poisson fordelingen:

Punktsannsynlighet: $P(X = k) = \frac{\lambda^k}{k!} e^{-\lambda} \quad k = 0, 1, \dots$

Momentgenererende funksjon : $M_X(t) = e^{\lambda(e^t - 1)}$

Forventning: $E(X) = \lambda$

Varians: $V(X) = \lambda$

Tilnærmelse: $Z = \frac{X - \lambda}{\sqrt{\lambda}}$ er tilnærmet standard normalfordelt
når λ er tilstrekkelig stor (minst 10)

Addisjonsregel: $X \sim \text{Poisson}(\lambda_1), Y \sim \text{Poisson}(\lambda_2)$
og X, Y uavhengige $\Rightarrow X + Y \sim \text{Poisson}(\lambda_1 + \lambda_2)$

e) Multinomisk fordeling:

Punktsannsynlighet: $P(N_1 = n_1, \dots, N_r = n_r) = \frac{n!}{n_1! \dots n_r!} p_1^{n_1} \dots p_r^{n_r}$

Her er $\sum_{i=1}^r p_i = 1$ og $\sum_{i=1}^r n_i = n$

Marginalfordeling: $N_i \sim \text{binomisk}(n, p_i)$

9. Noen kontinuerlige sannsynlighetsfordelinger

a) Normalfordelingen:

Tetthet: $f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/2\sigma^2} \quad -\infty < x < \infty$

Momentgenererende funksjon: $M_X(t) = e^{\mu t + \sigma^2 t^2/2}$

Forventning: $E(X) = \mu$

Varians: $V(X) = \sigma^2$

Transformasjon: $X \sim N(\mu, \sigma^2) \Rightarrow a + bX \sim N(a + b\mu, b^2\sigma^2)$

$X \sim N(\mu, \sigma^2) \Rightarrow Z = \frac{X-\mu}{\sigma} \sim N(0, 1)$

Addisjonsregel: $X \sim N(\mu_X, \sigma_X^2), Y \sim N(\mu_Y, \sigma_Y^2), X, Y$ uavhengige
 $\Rightarrow X + Y \sim N(\mu_X + \mu_Y, \sigma_X^2 + \sigma_Y^2)$

b) Eksponentialfordelingen:

Tetthet: $f(x) = \lambda e^{-\lambda x} \quad x > 0$

Momentgenererende funksjon: $M_X(t) = \lambda/(\lambda - t)$ for $t < \lambda$

Forventning: $E(X) = 1/\lambda$

Varians: $V(X) = 1/\lambda^2$

Addisjonsregel: $X \sim \exp(\lambda), Y \sim \exp(\lambda), X$ og Y uavhengige
 $\Rightarrow X + Y \sim \text{gamma}(2, 1/\lambda)$

c) Gammafordelingen:

Tetthet: $f(x) = \frac{1}{\beta^\alpha \Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} \quad x > 0$

Gammafunksjonen: $\Gamma(\alpha) = \int_0^\infty u^{\alpha-1} e^{-u} du$
 $\Gamma(\alpha + 1) = \alpha\Gamma(\alpha)$
 $\Gamma(n) = (n - 1)!$ når n er et helt tall
 $\Gamma(1/2) = \sqrt{\pi}$, $\Gamma(1) = 1$

Momentgenererende funksjon : $M_X(t) = [1/(1 - \beta t)]^\alpha$

Forventning: $E(X) = \alpha\beta$

Varians: $V(X) = \alpha\beta^2$

Addisjonsregel: $X \sim \text{gamma}(\alpha, \beta)$, $Y \sim \text{gamma}(\delta, \beta)$,
 X og Y uavhengige $\Rightarrow X + Y \sim \text{gamma}(\alpha + \delta, \beta)$

d) Kji-kvadratfordelingen:

Tetthet: $f(v) = \frac{1}{2^{n/2}\Gamma(n/2)} v^{(n/2)-1} e^{-v/2} \quad v > 0$

n er antall frihetsgrader

Forventning: $E(V) = n$

Varians: $V(V) = 2n$

Addisjonsregel: $U \sim \chi_n^2$, $V \sim \chi_m^2$, U og V uavhengige
 $\Rightarrow U + V \sim \chi_{n+m}^2$

Resultat: $Z \sim N(0, 1) \Rightarrow Z^2 \sim \chi_1^2$

e) Students t-fordeling:

Tetthet: $f(t) = \frac{\Gamma[(n+1)/2]}{\sqrt{n\pi}\Gamma(n/2)} (1 + \frac{t^2}{n})^{-(n+1)/2} \quad -\infty < t < \infty$

n er antall frihetsgrader

Forventning: $E(T) = 0 \quad (n \geq 2)$

Varians: $V(T) = n/(n - 2) \quad (n \geq 3)$

Resultat: $Z \sim N(0, 1)$, $U \sim \chi_n^2$, Z, U uavhengige $\Rightarrow Z/\sqrt{U/n} \sim t_n$

f) Binormal fordeling:

Tetthet:

$$f(x, y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \left[\frac{(x-\mu_X)^2}{\sigma_X^2} + \frac{(y-\mu_Y)^2}{\sigma_Y^2} - 2\rho \frac{(x-\mu_X)(y-\mu_Y)}{\sigma_X\sigma_Y} \right] \right\}$$

Marginalfordeling: $X \sim N(\mu_X, \sigma_X^2)$, $Y \sim N(\mu_Y, \sigma_Y^2)$

Korrelasjon: $\text{Corr}(X, Y) = \rho$

Betinget fordeling: Gitt $X = x$ er Y normalfordelt med
forventning $E(Y|X = x) = \mu_Y + \rho \frac{\sigma_Y}{\sigma_X} (x - \mu_X)$
og varians $V(Y|X = x) = \sigma_Y^2 (1 - \rho^2)$

10. Maksimum likelihood metoden

Anta at X_1, X_2, \dots, X_n har simultan punktsannsynlighet/sannsynlighetstetthet $f(x_1, x_2, \dots, x_n | \boldsymbol{\theta})$, der $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)$ er en parametervektor (skalar hvis $p = 1$). Vi antar at $f(x_1, x_2, \dots, x_n | \boldsymbol{\theta})$ tilfredsstiller visse deriverbarhetsbetingelser.

- Gitt observerte verdier $X_i = x_i$; $i = 1, \dots, n$; er likelihood-funksjonen $L(\boldsymbol{\theta}) = f(x_1, x_2, \dots, x_n | \boldsymbol{\theta})$ og loglikelihood-funksjonen $l(\boldsymbol{\theta}) = \log L(\boldsymbol{\theta})$.
- Maksimum likelihood *estimatet* er den verdien av $\boldsymbol{\theta}$ som maksimerer $L(\boldsymbol{\theta})$ eller ekvivalent maksimerer $l(\boldsymbol{\theta})$. Hvis vi erstatter de observerte x_i -ene med de stokastiske X_i -ene, får vi maksimum likelihood *estimatoren*.
- Maksimum likelihood estimatet $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_p)$ er en løsning av ligningene $s_j(\boldsymbol{\theta}) = 0$; $j = 1, \dots, p$; der $s_j(\boldsymbol{\theta}) = (\partial/\partial\theta_j)l(\boldsymbol{\theta})$ er score-funksjonene. Vektoren av scorefunksjoner er $\mathbf{s}(\boldsymbol{\theta}) = (s_1(\boldsymbol{\theta}), \dots, s_p(\boldsymbol{\theta}))^T$.
- Den observerte informasjonsmatrisen $\bar{\mathbf{J}}(\boldsymbol{\theta})$ er $p \times p$ matrisen med element (i, j) gitt ved $\bar{J}_{ij}(\boldsymbol{\theta}) = -\frac{\partial^2}{\partial\theta_i\partial\theta_j}l(\boldsymbol{\theta})$.

Den forventede informasjonsmatrisen (eller Fishers informasjonsmatrise) $\bar{\mathbf{I}}(\boldsymbol{\theta})$ er $p \times p$ matrisen med element (i, j) gitt ved $\bar{I}_{ij}(\boldsymbol{\theta}) = E[\bar{J}_{ij}(\boldsymbol{\theta})]$.

For uavhengige og identisk fordelte observasjoner har vi at $\bar{\mathbf{I}}(\boldsymbol{\theta}) = n\mathbf{I}(\boldsymbol{\theta})$ der $\mathbf{I}(\boldsymbol{\theta})$ er forventet informasjon til en observasjon.

- Når vi har "tilstrekkelig mye" data, er $\hat{\theta}_i$ tilnærmet normalfordelt med forventning θ_i og med varians lik det i -te diagonalelementet til $\bar{\mathbf{I}}^{-1}(\boldsymbol{\theta})$. Kovariansen mellom $\hat{\theta}_i$ og $\hat{\theta}_j$ er tilnærmet lik element (i, j) i $\bar{\mathbf{I}}^{-1}(\boldsymbol{\theta})$. Vi kan estimere varianser/kovarianser ved å sette inn $\hat{\boldsymbol{\theta}}$ for $\boldsymbol{\theta}$ i $\bar{\mathbf{I}}^{-1}(\boldsymbol{\theta})$ eller i $\bar{\mathbf{J}}^{-1}(\boldsymbol{\theta})$.
- Anta at vi ønsker å teste $H_0 : \theta_{q+1} = \dots = \theta_p = 0$ mot alternativet at minst en av $\theta_j \neq 0$, $j = q+1, \dots, p$. La $\hat{\boldsymbol{\theta}} = (\hat{\theta}_1, \dots, \hat{\theta}_p)$ være MLE i den generelle modellen og $\boldsymbol{\theta}^* = (\theta_1^*, \dots, \theta_q^*, 0, \dots, 0)$ MLE under nullhypotesen og dessuten $\nu = p - q$. Da er likelihood ratioen gitt som $LR = L(\boldsymbol{\theta}^*)/L(\hat{\boldsymbol{\theta}})$. Når nullhypotesen holder og vi har "tilstrekkelig mye" data gjelder tilnærmet at $-2 \log(LR) = 2[l(\hat{\boldsymbol{\theta}}) - l(\boldsymbol{\theta}^*)] \sim \chi_\nu^2$.

11. Ett normalfordelt utvalg

Hvis X_1, X_2, \dots, X_n er uavhengige og $N(\mu, \sigma^2)$ -fordelte så har vi at:

- $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ og $S^2 = \frac{1}{n-1} \sum_{i=1}^n (X_i - \bar{X})^2$ er uavhengige
- $\bar{X} \sim N(\mu, \sigma^2/n)$
- $(n-1)S^2/\sigma^2 \sim \chi_{n-1}^2$
- $\frac{\bar{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$

12. To normalfordelte utvalg

La X_1, X_2, \dots, X_n være uavhengige og $N(\mu_X, \sigma^2)$ -fordelte, og Y_1, Y_2, \dots, Y_m uavhengige og $N(\mu_Y, \sigma^2)$ -fordelte. De to utvalgene er uavhengige av hverandre. La \bar{X}, \bar{Y}, S_X^2 og S_Y^2 være definert i henhold til 11a). Da har vi at:

- $S_p^2 = [(n-1)S_X^2 + (m-1)S_Y^2]/(m+n-2)$ er en vektet estimator for σ^2
- $\bar{X} - \bar{Y} \sim N(\mu_X - \mu_Y, \sigma^2(\frac{1}{n} + \frac{1}{m}))$
- $(n+m-2)S_p^2/\sigma^2 \sim \chi_{m+n-2}^2$
- $\frac{\bar{X} - \bar{Y} - (\mu_X - \mu_Y)}{S_p \sqrt{\frac{1}{n} + \frac{1}{m}}} \sim t_{m+n-2}$

13. Regresjonsanalyse

Anta at $Y_i = \beta_0 + \beta_1 x_i + \epsilon_i$; $i = 1, 2, \dots, n$; hvor x_i -ene er kjente tall og ϵ_i -ene er uavhengige og $N(0, \sigma^2)$ -fordelte. Da har vi at:

- Minste kvadraters estimatorer for β_0 og β_1 er

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x} \quad \text{og} \quad \hat{\beta}_1 = \frac{\sum_{i=1}^n (x_i - \bar{x})(Y_i - \bar{Y})}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

- Estimatorene i a) er normalfordelte og forventningsrette, og

$$\text{Var}(\hat{\beta}_0) = \frac{\sigma^2 \sum_{i=1}^n x_i^2}{n \sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{og} \quad \text{Var}(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

- La $\text{SSE} = \sum_{i=1}^n (Y_i - \hat{\beta}_0 - \hat{\beta}_1 x_i)^2$. Da er $S^2 = \text{SSE}/(n-2)$ en forventningsrett estimator for σ^2 , og $(n-2)S^2/\sigma^2 \sim \chi_{n-2}^2$

14. Multippel lineær regresjon

Anta at $Y_i = \beta_0 + \beta_1 x_{i1} + \dots + \beta_k x_{ik} + \epsilon_i$; $i = 1, 2, \dots, n$; der x_{ij} -ene er kjente tall og ϵ_i -ene er uavhengige og $N(0, \sigma^2)$ -fordelte. På matriseform kan vi skrive modellen som $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta}$, der $\mathbf{Y} = (Y_1, \dots, Y_n)^T$ og $\boldsymbol{\beta} = (\beta_0, \dots, \beta_k)^T$ er henholdsvis n - og $(k+1)$ -dimensjonale vektorer, og $\mathbf{X} = \{x_{ij}\}$ (med $x_{i0} = 1$) er en $n \times (k+1)$ -dimensjonal matrise. Vi har at:

- Minste kvadraters estimator for $\boldsymbol{\beta}$ er $\hat{\boldsymbol{\beta}} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{Y}$.

- La $\hat{\boldsymbol{\beta}} = (\hat{\beta}_0, \dots, \hat{\beta}_k)^T$. Da er $\hat{\beta}_j$ -ene normalfordelte og forventningsrette, og

$$\text{Var}(\hat{\beta}_j) = \sigma^2 c_{jj} \quad \text{og} \quad \text{Cov}(\hat{\beta}_j, \hat{\beta}_l) = \sigma^2 c_{jl}$$

der c_{jl} er element (j, l) i $(k+1) \times (k+1)$ matrisen $\mathbf{C} = (\mathbf{X}^T \mathbf{X})^{-1}$.

- c) La $\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 x_{i1} + \dots + \hat{\beta}_k x_{ik}$, og sett $SSE = \sum_{i=1}^n (Y_i - \hat{Y}_i)^2$. Da er $S^2 = SSE/[n - (k + 1)]$ en forventningsrett estimator for σ^2 , og $[n - (k + 1)]S^2/\sigma^2 \sim \chi_{n-(k+1)}^2$. Videre er S^2 og $\hat{\beta}$ uavhengige.
- e) La $S_{\hat{\beta}_j}^2$ være den variansestimatorene for $\hat{\beta}_j$ vi får ved å erstatte σ^2 med S^2 i formelen for $\text{Var}(\hat{\beta}_j)$ i punkt b). Da er $(\hat{\beta}_j - \beta_j)/S_{\hat{\beta}_j} \sim t_{n-(k+1)}$.