FORMELSAMLING TIL STK1110

(Versjon av 23. november 2022)

1. Sannsynlighet

La $A, B, A_1, A_2, \ldots, B_1, B_2, \ldots$ være begivenheter, dvs. delmengder av et utfallsrom $\mathcal{S}$

a) Aksiomene:
Et sannsynlighetsmål P er en funksjon fra delmengder av utfallsrommet $\mathcal{S}$ til de reelle tall som tilfredsstiller

$$P(A) \geq 0$$
$$P(\mathcal{S}) = 1$$
$$P\Big(\bigcup_{i=1}^{\infty} A_i\Big) = \sum_{i=1}^{\infty} P(A_i) \quad \text{hvis } A_i \cap A_j = \emptyset \text{ for } i \neq j$$

b) $P(\emptyset) = 0$

c) $P\Big(\bigcup_{i=1}^{k} A_i\Big) = \sum_{i=1}^{k} P(A_i)$ hvis $A_i \cap A_j = \emptyset$ for $i \neq j$

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

e) $P(A) \leq 1$

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

f) Betinget sannsynlighet:

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

g) Total sannsynlighet:

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

h) Bayes' setning:

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

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

j) $A_1, \ldots, A_n$ er uavhengige begivenheter dersom

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

for alle delmengder av indekser $i_1, i_2, \ldots, i_k$

k) Produktsetningen:

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

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

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

3. Sannsynlighetsfordelinger

a) For en stokastisk variabel X (diskret eller kontinuerlig) er den kumulative fordelingsfunksjonen $F(x) = P(X \leq x)$

b) For en diskret stokastisk variabel X som kan anta verdiene $x_1, x_2, x_3, \ldots$ 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 kontinuerlig stokastisk variabel X har vi

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

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 kontinuerlige) 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 $x_1, x_2, \ldots$ og $y_1, y_2, \ldots$ 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

$$P\,(\,(X,Y) \in A\,) = \int\int_A f(u,v)dvdu$$
$$F(x,y) = \int_{-\infty}^{x} \int_{-\infty}^{y} f(u,v)dvdu$$
$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y}$$

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

g) Marginale punktsannsynligheter:

$$p_X(x_i) = \sum_j p(x_i, y_j) \qquad \text{(for } X\text{)}$$
$$p_Y(y_j) = \sum_i p(x_i, y_j) \qquad \text{(for } Y\text{)}$$

h) Marginale sannsynlighetstettheter:

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

i) Uavhengighet:

De stokastiske variablene X og Y er uavhengige dersom

$$p(x_i, y_j) = p_X(x_i)p_Y(y_j) \qquad \text{(diskret)}$$
$$f(x, y) = f_X(x)f_Y(y) \qquad \text{(kontinuerlig)}$$

j) Betingete punktsannsynligheter:

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

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:

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

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

$$\mathrm{E}(X) = \sum_j x_j p(x_j) \qquad \text{(diskret)}$$
$$\mathrm{E}(X) = \int_{-\infty}^{\infty} x f(x) dx \qquad \text{(kontinuerlig)}$$

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

$$\mathrm{E}[g(X)] = \sum_j g(x_j)p(x_j) \qquad \text{(diskret)}$$
$$\mathrm{E}[g(X)] = \int_{-\infty}^{\infty} g(x)f(x)dx \qquad \text{(kontinuerlig)}$$

c) $\mathrm{E}(a + bX) = a + b\mathrm{E}(X)$

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

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

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

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

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

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

h) Betinget forventning:

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

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

5. Varians og standardavvik

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

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

b) $\mathrm{V}(X) = \mathrm{E}(X^2) - \big(\mathrm{E}(X)\big)^2$

c) $\mathrm{V}(a + bX) = b^2\,\mathrm{V}(X)$

d) Hvis $X_1, \ldots, X_n$ er uavhengige har vi

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

e)

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

f) Chebyshevs ulikhet:

La X være en stokastisk variabel med $\mu = \mathrm{E}(X)$ og $\sigma^2 = \mathrm{V}(X)$.
For alle $k > 0$ har vi

$$P(|X - \mu| > k\sigma) \leq \frac{1}{k^2}$$

6. Kovarians og korrelasjon

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

$$\mathrm{Cov}(X,Y) = \mathrm{E}\big[(X-\mu_X)(Y-\mu_Y)\big]$$
$$\rho = \mathrm{Corr}(X,Y) = \frac{\mathrm{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

b) $\mathrm{Cov}(X,X) = \mathrm{V}(X)$

c) $\mathrm{Cov}(X,Y) = \mathrm{E}(XY) - \mathrm{E}(X)\mathrm{E}(Y)$

d) X, Y uavhengige $\Rightarrow \mathrm{Cov}(X,Y) = 0$

e)

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

f) $-1 \le \mathrm{Corr}(X,Y) \le 1$ og $\mathrm{Corr}(X,Y) = \pm 1$ 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 $M_X(t) = \mathrm{E}(e^{tX})$

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 $\mathrm{E}(X^r) = M_X^{(r)}(0)$

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

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

Forventning: $\mathrm{E}(X) = np$

Varians : $\mathrm{V}(X) = np(1-p)$

Tilnærmelse 1: $Z = \dfrac{X - np}{\sqrt{np(1-p)}}$ 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: $X \sim$ binomisk (n, p), $Y \sim$binomisk (m, p)
og X, Y uavhengige $\Rightarrow X + Y \sim$ binomisk $(n + m, p)$

b) Geometrisk fordeling:

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

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

Forventning: $\mathrm{E}(X) = 1/p$

Varians: $\mathrm{V}(x) = (1-p)/p^2$

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

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

Forventning: $\mathrm{E}(X) = r(1-p)/p$

Varians: $\mathrm{V}(X) = r(1-p)/p^2$

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

d) Hypergeometrisk fordeling:

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

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

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

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

e) Poisson-fordelingen:

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

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

Forventning: $\mathrm{E}(X) = \lambda$

Varians: $\mathrm{V}(X) = \lambda$

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

Addisjonsregel: $X \sim$Poisson (λ_1), $Y \sim$Poisson (λ_2) og X, Y uavhengige $\Rightarrow X + Y \sim$ Poisson $(\lambda_1 + \lambda_2)$

e) Multinomisk fordeling:

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

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

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

9. Noen kontinuerlige sannsynlighetsfordelinger

a) Normalfordelingen:

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

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

Forventning: $\mathrm{E}(X) = \mu$

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

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

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

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

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

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

c) Gammafordelingen:

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

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
$\Gamma(1/2) = \sqrt{\pi}, \quad \Gamma(1) = 1$

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

Forventning: $\mathrm{E}(X) = \alpha\beta$

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

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

d) Kji-kvadratfordelingen:

Tetthet: $f(u) = \frac{1}{2^{n/2}\Gamma(n/2)} u^{(n/2)-1} e^{-u/2} \qquad u > 0$
n er antall frihetsgrader

Forventning: $\mathrm{E}(U) = n$

Varians: $\mathrm{V}(U) = 2n$

Addisjonsregel: $U \sim \chi^2_n$, $V \sim \chi^2_m$, U og V uavhengige
$\Rightarrow U + V \sim \chi^2_{n+m}$

Resultat: $Z \sim N(0,1) \Rightarrow Z^2 \sim \chi^2_1$

e) Students t-fordeling:

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

Forventning: $\mathrm{E}(T) = 0 \qquad (n \geq 2)$

Varians: $\mathrm{V}(T) = n/(n-2) \qquad (n \geq 3)$

Resultat: $Z \sim N(0,1), \quad U \sim \chi^2_n, \quad Z, U$ 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: $X \sim N(\mu_X, \sigma_X^2), \quad Y \sim N(\mu_Y, \sigma_Y^2)$

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

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

10. Transformasjoner av kontinuerlige stokastiske variabler

a) Anta at X har sannsynlighetstetthet $f(x)$ og la $Y = u(X)$, der u er deriverbar og strengt monoton (vosende eller avtagende). La v være den inverse funksjonen til u, slik at $X = v(Y)$. Da er sannsynlighetstettheten til Y gitt ved

$$g(y) = f(v(y))|v'(y)|$$

b) Anta at (X_1, X_2) har simultantetthet $f(x_1, x_2)$. La

$$(Y_1, Y_2) = [u_1(X_1, X_2), u_2(X_1, X_2)]$$

være en en-entydig transformasjon av X_i-ene, og uttrykk X_i-ene ved Y_i-ene som

$$(X_1, X_2) = [v_1(Y_1, Y_2), v_2(Y_1, Y_2)]$$

Da er den simultane tettheten til (Y_1, Y_2) gitt ved

$$g(y_1, y_2) = f(v_1(y_1, y_2), v_2(y_1, y_2)) \left| \frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} \right|$$

der

$$\frac{\partial(x_1, x_2)}{\partial(y_1, y_2)} = \frac{\partial x_1}{\partial y_1}\frac{\partial x_2}{\partial y_2} - \frac{\partial x_2}{\partial y_1}\frac{\partial x_1}{\partial y_2}$$

er Jacobi-determinanten

11. Momentmetoden

Anta at $X_1, \ldots, X_n$ er uavhengige stokastiske variabler med tetthet/punktsannsynlighet som avhenger av parameterene $\theta_1, \ldots, \theta_m$. Da er

$$\mu_k(\theta_1, \ldots, \theta_m) = E(X^k)$$

det k-te teoretiske momentet og

$$\frac{1}{n}\sum_{i=1}^{n} X_i^k$$

er det k-te empiriske momentet. Momentestimatene er de verdiene av $\theta_1, \ldots, \theta_m$ som gjør at de m første teoretiske og empiriske momentene blir like. Momentestimatene er dermed løsningen av likningene

$$\mu_k(\theta_1, \ldots, \theta_m) = \frac{1}{n}\sum_{i=1}^{n} X_i^k$$

for $k = 1, \ldots, m$.

12. Maksimum likelihood-metoden

Anta at $X_1, X_2, \ldots, X_n$ er uavhengige og identisk fordelte med tetthet/punktsannsynlighet $f(x \mid \theta)$ som avhenger av én parameter θ. Vi antar at $f(x \mid \theta)$ tilfredsstiller visse regularitetsbetingelser.

a) Gitt observerte verdier $X_i = x_i$; $i = 1, \ldots, n$; er likelihood-funksjonen $L(\theta) = \prod_{i=1}^{n} f(x_i \mid \theta)$ og log-likelihood-funksjonen $\ell(\theta) = \log L(\theta)$.

b) Maksimum likelihood-*estimatet* er den verdien av θ som maksimerer $L(\theta)$ eller ekvivalent maksimerer $\ell(\theta)$. Hvis vi erstatter de observerte x_i-ene med de stokastiske X_i-ene, får vi maksimum likelihood-*estimatoren*.

c) Maksimum likelihood-estimatet $\widehat{\theta}$ er løsning av ligningen $s(\theta) = 0$, der $s(\theta) = (\partial/\partial\theta)\ell(\theta)$ er score-funksjonen.

d) Fisher-informasjonen i én observasjon er

$$I(\theta) = \mathrm{V}\left(\frac{\partial}{\partial\theta} \ln f(X \mid \theta)\right) = -\mathrm{E}\left(\frac{\partial^2}{\partial\theta^2} \ln f(X \mid \theta)\right)$$

og informasjonen i hele utvalget er $I_n(\theta) = nI(\theta)$.

e) Når vi har "tilstrekkelig mye" data, er $\widehat{\theta}$ tilnærmet normalfordelt med forventning θ og varians $1/[nI(\theta)]$.

13. Ett normalfordelt utvalg

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

a) $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ og $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$ er uavhengige

b) $\overline{X} \sim N(\mu, \sigma^2/n)$

c) $(n-1)S^2/\sigma^2 \sim \chi^2_{n-1}$

d) $\frac{\overline{X}-\mu}{S/\sqrt{n}} \sim t_{n-1}$

14. To normalfordelte utvalg

La $X_1, X_2, \ldots, X_m$ være uavhengige og $N(\mu_1, \sigma^2)$-fordelte, og $Y_1, Y_2, \ldots, Y_n$ uavhengige og $N(\mu_2, \sigma^2)$-fordelte. De to utvalgene er uavhengige av hverandre. La $\overline{X}$ og S_1^2 være definert i henhold til 13a) for X_i-ene og la $\overline{Y}$ og S_2^2 være definert tilsvarende for Y_i-ene. Da har vi at:

a) $S_p^2 = [(m-1)S_1^2 + (n-1)S_2^2]/(m+n-2)$ er en vektet estimator for σ^2

b) $\overline{X} - \overline{Y} \sim N\left(\mu_1 - \mu_2, \sigma^2(\frac{1}{n} + \frac{1}{m})\right)$

c) $(m+n-2)S_p^2/\sigma^2 \sim \chi^2_{m+n-2}$

d) $\frac{\overline{X}-\overline{Y}-(\mu_1-\mu_2)}{S_p\sqrt{\frac{1}{m}+\frac{1}{n}}} \sim t_{m+n-2}$

15. Enveis variansanalyse

Anta at $X_{ij} = \mu + \alpha_i + \epsilon_{ij}$; $j = 1, 2, \ldots, J$; $i = 1, 2, \ldots, I$; der ϵ_{ij}-ene er uavhengige og $N(0, \sigma^2)$-fordelte og $\sum_{i=1}^{I} \alpha_i = 0$. Da har vi at:

a) Den totale kvadratsummen $\mathrm{SST} = \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \overline{X}_{\cdot\cdot})^2$ kan skrives som SST = SSE + SSTr, der

$\mathrm{SSE} = \sum_{i=1}^{I}\sum_{j=1}^{J}(X_{ij} - \overline{X}_{i\cdot})^2$ er kvadratsummen for feil eller kvadratsummen innen ("within") grupper,

$\mathrm{SSTr} = J\sum_{i=1}^{I}(\overline{X}_{i\cdot} - \overline{X}_{\cdot\cdot})^2$ er kvadratsummen for behandling eller kvadratsummen mellom ("between") grupper.

b) SSE og SSTr er uavhengige.

c) $\mathrm{MSE} = \mathrm{SSE}/[I(J-1)]$ er en forventningsrett estimator for σ^2.
SSE/σ^2 er kji-kvadratfordelt med $I(J-1)$ frihetsgrader.

d) Hvis alle α_i-ene er lik null, er SSTr/σ^2 kji-kvadratfordelt med $I-1$ frihetsgrader.

e) Hvis alle α_i-ene er lik null, er $F = \frac{\mathrm{SSTr}/(I-1)}{\mathrm{SSE}/[I(J-1)]}$ F-fordelt med $I-1$ og $I(J-1)$ frihetsgrader

16. Regresjonsanalyse

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

a) Minste kvadraters estimatorer for β_0 og β_1 er

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

b) Estimatorene i a) er normalfordelte og forventningsrette, og

$$\mathrm{V}(\widehat{\beta}_0) = \frac{\sigma^2\sum_{i=1}^{n} x_i^2}{n\sum_{i=1}^{n}(x_i - \overline{x})^2} \quad \text{og} \quad \mathrm{V}(\widehat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^{n}(x_i - \overline{x})^2}$$

c) La $\mathrm{SSE} = \sum\limits_{i=1}^{n}(Y_i - \widehat{\beta}_0 - \widehat{\beta}_1 x_i)^2$. Da er $S^2 = \mathrm{SSE}/(n-2)$ en forventningsrett estimator for σ^2, og $(n-2)S^2/\sigma^2 \sim \chi^2_{n-2}$

17. Multippel lineær regresjon

Anta at $Y_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_k x_{ik} + \epsilon_i\,;\; i = 1, 2, \ldots, n\,;$ 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 $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\epsilon}$, der $\mathbf{Y} = (Y_1, \ldots, Y_n)'$, $\boldsymbol{\epsilon} = (\epsilon_1, \ldots, \epsilon_n)'$ og $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)'$ er henholdsvis n-, 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:

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

b) La $\widehat{\boldsymbol{\beta}} = (\widehat{\beta}_0, \ldots, \widehat{\beta}_k)'$. Da er $\widehat{\beta}_j$-ene normalfordelte og forventningsrette, og

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

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

c) La $\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_{i1} + \cdots + \widehat{\beta}_k x_{ik}$, og sett $SSE = \sum\limits_{i=1}^{n}(Y_i - \widehat{Y}_i)^2$. Da er $S^2 = \mathrm{SSE}/[n-(k+1)]$ en forventningsrett estimator for σ^2 og $[n-(k+1)]S^2/\sigma^2 \sim \chi^2_{n-(k+1)}$. Videre er S^2 og $\widehat{\boldsymbol{\beta}}$ uavhengige.

e) La $S^2_{\widehat{\beta}_j}$ være den variansestimatoren for $\widehat{\beta}_j$ vi får ved å erstatte σ^2 med S^2 i formelen for $\mathrm{V}(\widehat{\beta}_j)$ i punkt b). Da er $(\widehat{\beta}_j - \beta_j)/S_{\widehat{\beta}_j} \sim t_{n-(k+1)}$.

18. Sign-test og -konfidensintervall

Anta at $x_1, \ldots, x_n$ er observasjoner av de stokastiske variablene $X_1, \ldots, X_n$, som er uavhengig identisk fordelt fra en kontinuerlig fordeling.

a) La $Y_i = X_{(i)}$, der $X_{(1)} \leq \ldots \leq X_{(n)}$. Et $(1-\alpha) \cdot 100\%$ konfidensintervall for $100p\%$-persentilen η_p $(0 < p < 1)$ oppnås da ved å finne r og s slik at $\mathrm{P}(Y_r \leq \eta_p \leq Y_s) = \sum_{k=r}^{s-1}\binom{n}{k}p^k(1-p)^{n-k} \geq 1-\alpha$ kommer nærmest mulig $1-\alpha$.

b) For hypotesetester vedrørende medianen $\tilde{\mu}$ til X_i, omformulér til hypotesetest vedrørende $p = \mathrm{P}(X_i > \tilde{\mu}_0)$ med testobservator $T = \sum_{i=1}^{n} I(X_i > \tilde{\mu}_0)$, der $T \sim Bin(n, 0.5)$ når $p = 0.5$:

- $H_0 : \tilde{\mu} \leq \tilde{\mu}_0$ mot $H_a : \tilde{\mu} > \tilde{\mu}_0$:
  test $H_0 : p \leq 0.5$ mot $H_a : p > 0.5$.
- $H_0 : \tilde{\mu} \geq \tilde{\mu}_0$ mot $H_a : \tilde{\mu} < \tilde{\mu}_0$:
  test $H_0 : p \geq 0.5$ mot $H_a : p < 0.5$.
- $H_0 : \tilde{\mu} = \tilde{\mu}_0$ mot $H_a : \tilde{\mu} \neq \tilde{\mu}_0$:
  test $H_0 : p = 0.5$ mot $H_a : p \neq 0.5$.

19. Wilcoxon signed rank-test og -konfidensintervall

Anta at $x_1, \ldots, x_n$ er observasjoner av de stokastiske variablene $X_1, \ldots, X_n$, som er uavhengig identisk fordelt fra en symmetrisk, kontinuerlig fordeling

a) Anta at vi vil teste hypoteser knyttet til medianen $\tilde{\mu}$ til X_i, med $\tilde{\mu}_0$ som nullverdi i testen, dvs. en av
$H_0 : \tilde{\mu} \leq \tilde{\mu}_0$ mot $H_a : \tilde{\mu} > \tilde{\mu}_0$
$H_0 : \tilde{\mu} \geq \tilde{\mu}_0$ mot $H_a : \tilde{\mu} < \tilde{\mu}_0$
$H_0 : \tilde{\mu} = \tilde{\mu}_0$ mot $H_a : \tilde{\mu} \neq \tilde{\mu}_0$.
La så $R_1, \ldots, R_n$ være rangene til $|X_1 - \tilde{\mu}_0|$,...,$|X_n - \tilde{\mu}_0|$. Da bruker vi testobservatoren $S_+ = \sum_{i=1}^n I(X_i - \tilde{\mu}_0 > 0) \cdot R_i$.

b) For store utvalg ($n > 20$), er $S_+ \overset{tiln.}{\sim} N\left(\frac{n(n+1)}{4}, \frac{n(n+1)(2n+1)}{24}\right)$ når $\tilde{\mu} = \tilde{\mu}_0$.

c) For å finne et $(1 - \alpha) \cdot 100\%$ konfidensintervall for $\tilde{\mu}$, beregn Walsh-snittene $\overline{x}_1$,...,$\overline{x}_{n(n+1)/2}$, som $(x_i{+}x_j)/2$, $i = 1, \ldots, n$, $j = 1, \ldots, i$, og sortér dem i voksende rekkefølge $\overline{x}_{(1)} \leq \ldots \leq \overline{x}_{(n(n+1)/2)}$. Et $(1 - \alpha) \cdot 100\%$ konfidensintervall for $\tilde{\mu}$ gitt ved

$$\left(\overline{x}_{(n(n+1)/2)-c+1}, \overline{x}_{(c)}\right),$$

der c er slik at $H_0 : \tilde{\mu} = \tilde{\mu}_0$ forkastes til fordel for $H_a : \tilde{\mu} \neq \tilde{\mu}_0$ ved signifikansnivå α dersom $S_+ \geq c$ eller $S_+ \leq n(n + 1)/2 - c$.

20. Bayesiansk inferens

Anta at $x_1, \ldots, x_n$ er observasjoner av de stokastiske variablene $X_1, \ldots, X_n$, som er uavhengig identisk fordelt fra en fordeling med tetthet/punktsannsynlighet $f(x|\theta)$, som avhenger av én parameter θ. Videre antar vi at apriorifordelingen til θ har tetthet/punktsannsynlighet $\pi(\theta)$.

a) Aposteriorifordelingen til θ, altså fordelingen til $[\theta|X_1 = x_1, \ldots, X_n = x_n]$, har da tetthet/punktsannsynlighet gitt ved

$$\pi(\theta|x_1, \ldots, x_n) = \begin{cases} \frac{\pi(\theta) f(x_1,\ldots,x_n|\theta)}{\int_{-\infty}^{\infty} \pi(\theta) f(x_1,\ldots,x_n|\theta) d\theta}, & \theta \text{ er kontinuerlig} \\ \frac{\pi(\theta) f(x_1,\ldots,x_n|\theta)}{\sum_\theta \pi(\theta) f(x_1,\ldots,x_n|\theta)}, & \theta \text{ er diskret} \end{cases}.$$

b) Bayes-estimatoren for θ er $\widehat{\theta} = \mathrm{E}(\theta|X_1, \ldots, X_n)$, altså forventningen til aposteriorifordelingen, med tilsvarende estimat $\widehat{\theta} = \mathrm{E}(\theta|x_1, \ldots, x_n)$.

c) Et $(1 - \alpha) \cdot 100\%$ kredibilitetsintervall for θ er gitt ved $(\eta_{\alpha/2}, \eta_{1-\alpha/2})$, der η_p er $100p\%$-persentilen til aposteriorifordelingen.