Mamo3100

Per Erik Grønvik **2025**

Vanlige tegn i Statistikk

$\sum$: Summasjon, brukes for å summere verdier.
$\prod$: Produkt, brukes for å multiplisere verdier.
$\bar{x}$: Gjennomsnitt, representerer gjennomsnittet av et datasett.
$\hat{\beta}$: Estimat, en estimert parameterverdi, ofte brukt i regresjon.
ε: Feilledd, tilfeldige feil eller støy i en modell.
n: Antall, antall observasjoner eller datapunkter.
σ^2: Varians, måler spredningen i et datasett.
σ: Standardavvik, kvadratroten av variansen.
$P(A)$: Sannsynlighet, sannsynligheten for hendelsen A.
$f(x)$: Tetthetsfunksjon, sannsynlighetstetthet for en kontinuerlig variabel.
$F(x)$: Fordelingsfunksjon, kumulativ sannsynlighet for en variabel.
μ: Forventningsverdi, populasjonsgjennomsnittet.
x_i: Observasjon, den i-te observasjonen i et datasett.
X: Tilfeldig variabel, en stokastisk variabel.
z: Z-score, standardisert verdi for en observasjon.
$|x|$: Absoluttverdi, avstanden fra 0 på tallinjen.
': Transpnert. a' transponert av a
derfor orthogonal y = + , med $^\top \hat{y} = 0$ og $X^\top \hat{\varepsilon} = 0$

Regressjon

Sum of Squared Errors (SSE): Linear regression formula
$(x_1, y_1), \ldots, (x_n, y_n)$
$y = \alpha + \beta x + \varepsilon_i$
$\hat{\alpha} = \bar{y} - \hat{\beta}\bar{x}$
$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x}) y_i}{\sum_{i=1}^n (x_i - \bar{x})^2} = \frac{\sum_{i=1}^n X_i Y_i - n\bar{X}\bar{Y}}{\sum_{i=1}^n X_i^2 - n\bar{X}^2}$
$SSE(\alpha, \beta) = \sum (y_i - \alpha - \beta x_i)^2$
$\mathrm{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$
$\mathrm{Var}(\hat{\beta}) = \frac{\mathrm{Var}\left(\sum_{i=1}^n (X_i - \bar{X})\varepsilon_i\right)}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2} =$
$\frac{\sigma^2 \sum_{i=1}^n (X_i - \bar{X})^2}{\left(\sum_{i=1}^n (X_i - \bar{X})^2\right)^2} = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}.$

Konfidansintervall for T-interverall

$$\beta \pm t_{(\alpha, n-2)} \frac{\hat{\sigma}}{\sqrt{\sum_{i=1}^n (X_i - \overline{X})^2}}$$

Least Squares (LS):

$$\mathrm{LS}(\boldsymbol{\beta}) = \sum_{i=1}^n (y_i - \mathbf{x}_i' \boldsymbol{\beta})^2 = \sum_{i=1}^n \varepsilon_i^2 = \boldsymbol{\varepsilon}' \boldsymbol{\varepsilon},$$

Least Absolute Deviations (LA):

$$\mathrm{LA}(\boldsymbol{\beta}) = \sum_{i=1}^n |y_i - \mathbf{x}_i' \boldsymbol{\beta}| = \sum_{i=1}^n |\varepsilon_i|.$$

Rules for summation
$\sum_{i=1}^n (a + b) = \sum_{i=1}^n a + \sum_{i=1}^n b$
$\sum_{i=1}^n c \cdot a_i = c \cdot \sum_{i=1}^n a_i, \quad \text{if } c \text{ is constant.}$
$\sum_{i=1}^n c = n \cdot c, \quad \text{if } c \text{ is constant.}$
$\sum_{i=1}^n (a_i \cdot b_i) \neq \left(\sum_{i=1}^n a_i\right) \cdot \left(\sum_{i=1}^n b_i\right), \quad \text{(not always equal!)}$
$\sum_{i=1}^n i = \frac{n(n+1)}{2}$
$\sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}$
$\sum_{i=1}^n (X_i - \bar{X}) X_i = \sum_{i=1}^n (X_i - \bar{X})^2$
$\sum_{i=1}^n (X_i - \bar{X})^2 = \sum_{i=1}^n X_i^2 - n\bar{X}^2$
$\sum_{i=1}^n (aX_i + b)^2 = a^2 \sum_{i=1}^n X_i^2 + 2ab \sum_{i=1}^n X_i + nb^2$
$\sum_{i=1}^n (x_i - \bar{x}) = 0$ $\bar{X} = \frac{1}{n} \sum_{i=1}^n X_i$ $\mathrm{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (X_i - \bar{X})(Y_i - \bar{Y})$
$\mathrm{Var}(X) = E(X^2) - E(X)^2$

Linearity of Differentiation:

$$\frac{\partial}{\partial x} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{\partial f_n(x)}{\partial x},$$

provided the following conditions are satisfied:

- The series $\sum_{n=1}^{\infty} f_n(x)$ **converges uniformly** in the region where differentiation is applied.
- Each term $f_n(x)$ is **differentiable** in the region.

Rules for expectation E
$E(c) = c, \quad \text{if } c \text{ is constant.}$
$E(a + b) = E(a) + E(b)$
$E(c \cdot a) = c \cdot E(a), \quad \text{if } c \text{ is constant.}$
$E(a \cdot b) = E(a) \cdot E(b), \quad \text{if } a \text{ and } b \text{ are independent.}$
$E\left(\sum_{i=1}^n a_i\right) = \sum_{i=1}^n E(a_i)$
$E(a^2) = \mathrm{Var}(a) + E(a)^2$
$E\left(\frac{a}{b}\right) \neq \frac{E(a)}{E(b)}, \quad \text{(not always equal, except in special cases).}$
$E(g(a)) = \int_{-\infty}^{\infty} g(x) f_a(x)\, dx, \quad$ (for the continuous case with density fu
$E(g(a)) = \sum_x g(x) P(a = x), \quad \text{(for the discrete case).}$

Regneregler for varians

$$\mathrm{Var}\left(\sum_{i=1}^n a_i \varepsilon_i\right) = \sum_{i=1}^n a_i^2 \mathrm{Var}(\varepsilon_i),$$

Formler og regneregler for lineær regresjon

1. Modell og grunnleggende formler

Lineær regresjonsmodell: $Y_i = \alpha + \beta X_i + \varepsilon_i$ Forventning og varians:

$$E(Y_i) = \alpha + \beta X_i, \quad \mathrm{Var}(Y_i) = \sigma^2$$

2. Minste kvadraters metode (OLS)

Estimat for β:

$$\hat{\beta} = \frac{\sum_{i=1}^n (X_i - \bar{X}) Y_i}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Estimat for α:

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

Sum of Squared Errors (SSE):

$$SSE = \sum_{i=1}^n \left(Y_i - \hat{Y}_i\right)^2 = \sum_{i=1}^n \left(Y_i - \hat{\alpha} - \hat{\beta} X_i\right)^2$$

3. Varians og standardfeil

Estimert varians til residualene (s^2):

$$\hat{\sigma}^2 = s^2 = \frac{1}{n-2} \sum_{i=1}^n \left(Y_i - \hat{Y}_i\right)^2$$

Varians til $\hat{\beta}$:

$$\mathrm{Var}(\hat{\beta}) = \frac{\sigma^2}{\sum_{i=1}^n (X_i - \bar{X})^2}$$

Standardfeil for $\hat{\beta}$:

$$SE(\hat{\beta}) = \sqrt{\frac{s^2}{\sum_{i=1}^n (X_i - \bar{X})^2}}$$

4. Hypotesetesting for β

Null- og alternativhypotese:

$$H_0 : \beta = 0 \quad \text{(ingen sammenheng)},$$

$$H_1 : \beta \neq 0 \quad \text{(sammenheng eksisterer)}$$

T-test-statistikk:

$$t_{\text{obs}} = \frac{\hat{\beta}}{SE(\hat{\beta})} = \frac{\hat{\beta}\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2}}{s}$$

Forkastningsregel:

$$\text{Forkast } H_0 \text{ hvis } |t_{\text{obs}}| > t_{\alpha/2,n-2}$$

5. Konfidensintervall for β

$$\hat{\beta} - t_{\alpha/2,n-2} \cdot SE(\hat{\beta}) < \beta < \hat{\beta} + t_{\alpha/2,n-2} \cdot SE(\hat{\beta})$$

6. Fordeling av estimatene

Fordeling av $\hat{\beta}$:

$$\hat{\beta} \sim N\left(\beta, \text{Var}(\hat{\beta})\right)$$

Fordeling av $\hat{\alpha}$:

$$\hat{\alpha} \sim N\left(\alpha, \text{Var}(\hat{\alpha})\right)$$

7. Bruk av T-fordeling

Når variansen σ^2 ikke er kjent:

$$t = \frac{\hat{\beta} - \beta}{SE(\hat{\beta})} \sim T_{n-2}$$

8. Summasjonsregler

Summasjonsregler for X_i:

$$\sum_{i=1}^{n}(X_i - \bar{X}) = 0, \quad \sum_{i=1}^{n}(X_i - \bar{X})^2 = \text{total varians i } X$$

9. Linearitet av forventning

Lineær kombinasjon av forventninger:

$$E(aX + b) = aE(X) + b$$

10. Varians av lineær kombinasjon

Varians av skalering:

$$\text{Var}(aX) = a^2\text{Var}(X)$$

Varians av summen av uavhengige variabler:

$$\text{Var}(X+Y) = \text{Var}(X) + \text{Var}(Y) \quad \text{(hvis } X \text{ og } Y \text{ er uavhengige)}$$

Left to do:

- Navngi på en fornuftig måte formlene hittil.

Mathematical Functions

Important Integrals

$$\int_{-\infty}^{\infty} e^{-x^2}\,dx = \sqrt{\pi}.$$

$$\int_{-\infty}^{\infty} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}}, \quad a > 0.$$

$$\int_{-\infty}^{\infty} x^2 e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}} \cdot \frac{1}{2a}, \quad a > 0.$$

$$\int_{-\infty}^{\infty} x^{2n} e^{-ax^2}\,dx = \sqrt{\frac{\pi}{a}} \cdot \frac{(2n-1)!!}{(2a)^n}, \quad n \in \mathbb{N}_0,\ a > 0,$$

$$\int_a^b uv'\,dx = \left[uv\right]_a^b - \int_a^b u'v\,dx$$

$$\int_a^b f(u)u'\,dx = \int_a^b f(u)\,\frac{du}{dx}dx = \int_{u(a)}^{u(b)} f(u)\,du$$

$$u = ax + b \Rightarrow \quad \frac{du}{dx} = a \Rightarrow \quad du = a\,dx \Rightarrow \quad dx = \frac{du}{a}$$

Even an odd functions

$$f(-x) = f(x) \Rightarrow \text{even} \Rightarrow \int_{-a}^{a} f(x)\,dx = 2\int_0^a f(x)\,dx$$

$$f(-x) = -f(x) \Rightarrow \text{odd} \Rightarrow \int_{-a}^{a} f(x)\,dx = 0$$

$$else \Rightarrow neither$$

Important derivations and rules

$$\frac{d}{dx} f\big(g(x)\big) = f'\big(g(x)\big) \cdot g'(x).$$

$$\frac{d}{dx}\big(C \cdot f(x)\big) = C \cdot \frac{d}{dx} f(x),$$

Trigonometric functions

$$\int \cos(kx)dx = \frac{1}{k}\sin(kx) + C$$
$$\int \sin(kx)dx = -\frac{1}{k}\cos(kx) + C$$
$$\int \tan(x)dx = -ln\,|\cos(x)| + C$$
$$\int (1+\tan^2(x))dx = \tan(x) + C$$
$$\int \frac{1}{\cos^2(x)}dx = \tan(x) + C$$
$$\int \sin^2(x)\,dx = -\frac{1}{4}\sin(2x) - \frac{x}{2} + C.$$
$$\int \cos^2(x)dx = \frac{1}{4}\sin(2x) + \frac{x}{2} + C$$
$$\int \tan^2(x)dx = \tan(x) - x + C$$

Sine and Cosine Powers and Products:

$$\sin^3(x) = \frac{3}{4}\sin(x) - \frac{1}{4}\sin(3x)$$
$$\sin(3x) = 3\sin(x) - 4\sin^3(x)$$
$$\cos(3x) = 4\cos^3(x) - 3\cos(x)$$
$$\sin^2(x) = \frac{1}{2}(1-\cos(2x))$$
$$\cos^2(x) = \frac{1}{2}(1+\cos(2x))$$

$$\cos(2x) = \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1 = 1 - 2\sin^2(x)$$

$$\sin(2x) = 2\sin(x)\cos(x)$$
$$\tan(2x) = \frac{2\tan(x)}{1-\tan^2(x)}$$

Cotangent Definition:

$$\cot(\theta) = \frac{\cos(\theta)}{\sin(\theta)}$$

Addition formulas for cosine:

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y)$$
$$\cos(x-y) = \cos(x)\cos(y) + \sin(x)\sin(y)$$

Partial Derivatives of a Function $f(x,y)$:

$$\left.\frac{\partial f}{\partial x}\right|_{x=x_0,y=y_0} = \lim_{h\to 0}\frac{f(x_0+h,y_0) - f(x_0,y_0)}{h}$$
$$\left.\frac{\partial f}{\partial y}\right|_{x=x_0,y=y_0} = \lim_{h\to 0}\frac{f(x_0,y_0+h) - f(x_0,y_0)}{h}$$

Euler's Formulas for Cosine and Sine:

$$\cos(x) = \frac{e^{ix}+e^{-ix}}{2}, \quad \sin(x) = \frac{e^{ix}-e^{-ix}}{2i}$$

Euler's Formula

$$e^{\pm ix} = \cos(x) \pm i\sin(x)$$

Specific Forms of Euler's Formula (see +- syntax)

$$e^{-ix} = \cos(x) - i\sin(x)$$
$$e^{ix} = \cos(x) + i\sin(x)$$

Rules for complex konjugates

$z = a+bi \quad \Rightarrow \quad \overline{z} = a - bi$
$\overline{z_1+z_2} \quad \Rightarrow \quad \overline{z_1} + \overline{z_2}$
$\overline{z_1-z_2} \quad \Rightarrow \quad \overline{z_1} - \overline{z_2}$
$\overline{z_1 z_2} \quad \Rightarrow \quad \overline{z_1}\cdot\overline{z_2}$
$\overline{\frac{z_1}{z_2}} \quad \Rightarrow \quad \frac{\overline{z_1}}{\overline{z_2}}, \quad z_2 \neq 0$
$z \in \mathbb{R} \quad \Rightarrow \quad \overline{z} = z$
$\overline{i} \quad \Rightarrow \quad -i$
$\overline{e^z} \quad \Rightarrow \quad e^{\overline{z}}$
$|\overline{z}| \quad \Rightarrow \quad |z|$
$\overline{\overline{z}} \quad \Rightarrow \quad z$
$\overline{\langle\psi,\phi\rangle} \quad \Rightarrow \quad \langle\phi,\psi\rangle$

Imaginary numbers

$i^2 = -1$
$z = a + bi, \quad \text{where } a, b \in \mathbb{R}$
$|z| = \sqrt{a^2+b^2}$
$\overline{z} = a - bi$
$z\cdot\overline{z} = |z|^2 = a^2 + b^2$
$e^{i\theta} = \cos(\theta) + i\sin(\theta)$
$|e^{i\theta}| = 1$
$\left(e^{i\theta}\right)^2 = e^{i(2\theta)}$
$z^n = 1 \implies z_k = e^{i\frac{2\pi k}{n}}, \quad k = 0, 1, 2, \ldots, n-1$
$z_1 = r_1 e^{i\theta_1}, \quad z_2 = r_2 e^{i\theta_2} \implies z_1\cdot z_2 = r_1 r_2 e^{i(\theta_1+\theta_2)}$
$\frac{z_1}{z_2} = \frac{r_1}{r_2}e^{i(\theta_1-\theta_2)}$
$\left(e^{i\theta}\right)^n = e^{in\theta} = (\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta)$

Logarithmic Rules

$\log_b(xy) = \log_b(x) + \log_b(y)$
$\log_b\left(\frac{x}{y}\right) = \log_b(x) - \log_b(y)$
$\log_b(x^k) = k\cdot\log_b(x)$
$\log_b(1) = 0$
$\log_b(b) = 1$
$\log_b(x) = \frac{\ln(x)}{\ln(b)}$
$\ln(e^x) = x$
$\ln(1) = 0$
$\ln(ab) = \ln(a) + \ln(b)$
$\ln\left(\frac{a}{b}\right) = \ln(a) - \ln(b)$
$\ln(a^k) = k\cdot\ln(a)$

Rules for absolute value

$|ab| = |a|\cdot|b|$
$\left|\frac{a}{b}\right| = \frac{|a|}{|b|}, \quad b \neq 0$
$|a+b| \leq |a| + |b|$ (Trekantulikheten)
$||a| - |b|| \leq |a-b|$
$||a|| = |a|$
$|a| = 0 \quad \Leftrightarrow \quad a = 0$
$|a|^2 = a^2$
$a(x) \in \mathbb{R}_{\geq 0} \implies |a(x)| = a(x)$

Finite Difference Approximation:

$$f'(x_i) \approx \frac{f(x_{i+1}) - f(x_{i-1})}{2\Delta x}$$
$$f''(x_i) \approx \frac{f(x_{i+1}) - 2f(x_i) + f(x_{i-1})}{(\Delta x)^2}$$

The error is:

$$O(h^2)$$