more linear model theory

Recall: our model setup

True model:

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

$$i = 1, \dots, n$$

$$\varepsilon_i \overset{ind}{\sim} N(0, \sigma^2)$$

equivalent

$$y_i \overset{ind}{\sim} N(\beta_0 + \beta x_i, \sigma^2)$$

what assumptions are implicit in this?

1. Homoscedasticity

Every ε_i has SAME variance

2. Independence of noise/response

Last time showed

est $\hat{\beta}_0, \hat{\beta}_1$ via Least squares

or equivalently it's MLE

for $y_i \overset{ind}{\sim} N(\beta_0+\beta_1 x_i, \sigma^2)$

ASSUME NOTHING

only assume true model

is

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$$

assume $E(\varepsilon_i) = 0$

$\Rightarrow$ $\hat{\beta}_0, \hat{\beta}_1$ unbiased

or

$$E(\hat{\beta}_0) = \beta_0$$

$$E(\hat{\beta}_1) = \beta_1$$

also

$$\mathrm{var}(\hat{\beta}_1) = \frac{\sigma_\varepsilon^2}{S_{xx}}$$

$$S_{ab} = \sum_{i=1}^{n} (a_i - \bar{a})(b_i - \bar{b})$$

Let

$$\hat{\varepsilon}_i = y_i - \hat{y}_i$$ "the residual"

$$= y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)$$

$\hat{\varepsilon}_i$ are usually our estimator for ε_i

Assume homoscedasticity

$$\varepsilon_i \overset{ind}{\sim} [0, \sigma^2]$$

$$\hat{\sigma}^2_\varepsilon = \frac{\sum_{i=1}^{n} \hat{\varepsilon}_i^{\,2}}{n-2}$$

$$E(\hat{\sigma}^2_\varepsilon) = \sigma^2_\varepsilon$$

Now assume

$$\varepsilon_i \overset{iid}{\sim} N(0, \sigma^2)$$

then $$\frac{\hat{\beta}_{1,LS} - \beta_1}{\hat{\sigma}^2_\varepsilon / \sqrt{S_{xx}}} \sim t_{n-2}$$

How to check assumptions on noise? ε_i

Use residuals $\hat{\varepsilon}_i$

SOME VECTORS and MATRICES

I will try to underline vectors

Let $\underline{a}, \underline{b} \in \mathbb{R}^n$

By default all vectors are

column vectors ($n \times 1$ matrix)

$$\underline{a} = \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} \qquad \underline{b} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

$$\underline{a}^T \underline{b} = \begin{pmatrix} a_1 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

$$= a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$$

$$= \sum_{i=1}^{n} a_i b_i$$

$$\underline{a}^T \underline{a} = a_1^2 + a_2^2 + \cdots + a_n^2$$

$$\sqrt{\underline{a}^T\underline{a}} = \sqrt{a_1^2+a_2^2+\dots+a_n^2}$$

$$= \|\underline{a}\|_2$$

$$\underline{a}^T\underline{a} = \|\underline{a}\|_2^2$$

---

matrix multiplication

$$AB = C$$

$$(n\times m)\cdot(m\times p)$$

$$A \in \mathbb{R}^{n\times m}$$

$$B \in \mathbb{R}^{m\times p}$$

$$C \in \mathbb{R}^{n\times p}$$

Matrices and Random vectors

Let $\underline{y} \in \mathbb{R}^n$ be a random vector

$$E(\underline{y}) = \begin{pmatrix} E(y_1) \\ \vdots \\ E(y_n) \end{pmatrix}$$

$$\operatorname{cov}(y_i, y_j) = \sigma_{ij}$$

$$\sigma_{ij} = \sigma_{ji} \quad \sigma_{ii} = \operatorname{var}(y_i)$$

$$\operatorname{cov}(\underline{y}) = \{\sigma_{ij}\}$$

$$= \begin{pmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_{22} & & & \vdots \\ \vdots & & \ddots & & \\ \sigma_{n1} & \cdots & & & \sigma_{nn} \end{pmatrix}$$

<u>Thm A</u>

$\underline{y} \in \mathbb{R}^n$ random

$\underline{c} \in \mathbb{R}^n$ constant

$A \in \mathbb{R}^{m \times n}$ constant matrix if $\underline{z} = \underline{c} + A\underline{y}$

then

$$E(\underline{z}) = \underline{c} + \underline{A}\, E(\underline{y})$$

take the ith element of $\underline{z}$.

$$z_i = c_i + \sum_{j=1}^{n} a_{ij}\, y_j$$

$$E(z_i) = c_i + \sum_{j=1}^{n} a_{ij}\, E(y_j)$$

$$\Rightarrow E(\underline{z}) = \underline{c} + A\, E(\underline{y})$$

Theorem B

Same setup

if $\underline{Z} = \underline{c} + A\underline{Y}$

then $\text{cov}(\underline{Z}) = A\,\text{cov}(\underline{Y})A^T$

$\text{cov}(Z_i, Z_j)$

$$= \text{cov}\left(c_i + \sum_{k=1}^{n} a_{ik} Y_k, \; c_j + \sum_{l=1}^{n} a_{jl} Y_l\right)$$

$$= \text{cov}\left(\sum_{k=1}^{n} a_{ik} Y_k, \sum_{l=1}^{n} a_{jl} Y_l\right)$$

$$= \sum_{k=1}^{n} a_{ik} \left(\sum_{l=1}^{n} a_{jl} \operatorname{cov}(Y_k, Y_l) \right)$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} a_{jl} \operatorname{cov}(Y_k, Y_l)$$

$$= \sum_{k=1}^{n} \sum_{l=1}^{n} a_{ik} a_{jl} \sigma_{k,l} \quad \checkmark$$

Theorem C

$$E(\underline{X}^T A \underline{X}) = \operatorname{trace}(A\Sigma) + \underline{\mu}^T A \underline{\mu}$$

Thanks to Peng Ding for this proof!

$\underline{X}$ random vector with mean $\underline{\mu}$ and covariance Σ, A is fixed matrix

2 facts

$$\operatorname{cov}(\underline{Y}) = E(\underline{Y}\underline{Y}^T) - E(\underline{Y})E(\underline{Y}^T)$$

$$\Rightarrow$$

$$E(\underline{Y}\underline{Y}^T) = \operatorname{cov}(\underline{Y}) + E(\underline{Y})E(\underline{Y}^T)$$

$$E(\operatorname{trace}(W)) = \operatorname{trace}(E(W))$$

Pf.

$$E(\mathbf{X}^T A \mathbf{X}) = E(tr(\mathbf{X}^T A \mathbf{X}))$$

by properties of trace

$$= E(tr(A \mathbf{X}\mathbf{X}^T))$$

$$= tr(E(A \mathbf{X}\mathbf{X}^T))$$

$$= tr(A\, E(\mathbf{X}\mathbf{X}^T))$$

$$= tr(A(\Sigma + \boldsymbol{\mu}\boldsymbol{\mu}^T))$$

$$= tr(A\Sigma) + tr(A\boldsymbol{\mu}\boldsymbol{\mu}^T)$$

$= \operatorname{tr}(A\Sigma) + \operatorname{tr}(\underline{\mu}^T A \underline{\mu})$

$= \operatorname{tr}(A\Sigma) + \underline{\mu}^T A \underline{\mu}$

<u>Ch. 14</u> Q26.

Let Z_1, Z_2, Z_3, Z_4

be r.v. $\operatorname{var}(Z_i) = 1$

$\operatorname{cor}(Z_i, Z_j) = \rho \quad i \neq j$

Show

$Z_1 + Z_2 + Z_3 + Z_4$ uncor.

$Z_1 + Z_2 - Z_3 - Z_4$

Define

$$\underline{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

$$Az = z_1 + z_2 + z_3 + z_4$$

$$B = \begin{pmatrix} 1 & 1 & -1 & -1 \end{pmatrix}$$

$$Bz = z_1 + z_2 - z_3 - z_4$$

$$\Rightarrow \operatorname{cov}(Az, Bz)$$

Theorem D

$Y = AX \qquad Z = BX$

$\mathrm{cov}(Y, Z)$

$= \mathrm{cov}(AX, BX)$

$= A \,\mathrm{cov}(X)\, B^T$

$\mathrm{cov}(AZ, BZ)$

$= A \,\mathrm{cov}(Z)\, B^T$

$= \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} \sigma^2 & & & \rho \\ & \sigma^2 & & \\ & & \sigma^2 & \\ \rho & & & \sigma^2 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$

$= 0.$

Ch 14 Q30

$X_1, \dots, X_n$

$var(X_i) = \sigma^2$

$cov(X_i, X_j) = \rho\sigma^2$

$Var(\overline{X})$

$$\underline{X} = \begin{pmatrix} X_1 \\ \vdots \\ X_n \end{pmatrix} \quad \underline{\frac{1}{n}} = \begin{pmatrix} \frac{1}{n} \\ \vdots \\ \frac{1}{n} \end{pmatrix}$$

$$\operatorname{var}(\bar{X}) = \operatorname{cov}(\bar{x}, \bar{x})$$

$$= \operatorname{cov}\left(\frac{\underline{1}^T}{\underline{n}} \underline{X}, \frac{\underline{1}^T}{\underline{n}} \underline{X}\right)$$

use Theorem D

$$= \frac{\underline{1}^T}{\underline{n}} \operatorname{cov}(\underline{X}) \frac{\underline{1}}{\underline{n}}$$

$$= \frac{\underline{1}^T}{\underline{n}} \begin{pmatrix} \sigma^2 & & \rho\sigma^2 \\ & \ddots & \\ \rho\sigma^2 & & \sigma^2 \end{pmatrix} \frac{\underline{1}}{\underline{n}}$$

$$= \frac{\underline{1}^T}{\underline{n}} \begin{pmatrix} \frac{1}{n}\left((n-1)\rho\sigma^2 + \sigma^2\right) \\ \vdots \\ \frac{1}{n}\left((n-1)\rho\sigma^2 + \sigma^2\right) \end{pmatrix}$$

$$= \sum_{i=1}^{n} \frac{1}{n^2}\left((n-1)\rho\sigma^2 + \sigma^2\right)$$

$$= \frac{1}{n}\left((n-1)\rho\sigma^2 + \sigma^2\right)$$

if $\rho = 0$

$$\Rightarrow \operatorname{var}(\bar{x}) = \frac{\sigma^2}{n}$$

<u>Q31</u> <u>Ch 14</u>

$$\underline{z} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{pmatrix} \quad \operatorname{cov}(\underline{z}) = \sigma^2 I$$

$$U = Z_1 + Z_2 + Z_3 + Z_4$$

$$V = (Z_1 + Z_2) - (Z_3 + Z_4)$$

$$A = \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix}$$

$$B = \begin{pmatrix} 1 & 1 & -1 & -1 \end{pmatrix}$$

$$AZ = U \qquad AB = V$$

$$\operatorname{cov}(U, V)$$

$$= \operatorname{cov}(AZ, AB)$$

$$= A \operatorname{cov}(Z) B^T$$

$$= A \, \sigma^2 I \cdot B^T$$

$$= \sigma^2 A B^T$$

$$= \sigma^2 \begin{pmatrix} 1 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$= 0$$

Q25

$$y_{i1} \, , \, y_{i2} \quad , \, \bar{y}_i = \frac{(y_{i1} + y_{i2})}{2}$$

$$y_{i1} = \beta_0 + \beta_1 x_i + \varepsilon_{i1}$$

$$Y_{i2} = \beta_0 + \beta_1 X_i + \varepsilon_{i2}$$

$$\bar{Y}_i = \frac{Y_{i1} + Y_{i2}}{2}$$

$$= \frac{\beta_0 + \beta_1 X_i + \varepsilon_{i1} + \beta_0 + \beta_1 X_i + \varepsilon_{i2}}{2}$$

$$= \beta_0 + \beta_1 X_i + \frac{\varepsilon_{i1} + \varepsilon_{i2}}{2}$$

$$= \beta_0 + \beta_1 X_i + \varepsilon_i'$$

Solving this with least squares!