MSDS 596 Regression & Time Series

Lecture 01 Distributions of OLS parameter estimates

Department of Statistics
Rutgers University

Sept 26, 2022

Announcements

- Instructor: **Koulik Khamaru** (kk1241@rutgers.edu)
  Office hour: Monday 4-5pm, 403 Hill center

- HW1 due next Monday (Oct 3) in class.

Normal density function

- Let $Z \sim \mathcal{N}(\mu, \sigma^2)$ for some $-\infty < \mu < \infty$ and $\sigma^2 > 0$.
- $$f_{\mu,\sigma^2}(z) = \frac{1}{\sigma\sqrt{2\pi}} \cdot \exp\left\{-\frac{1}{2\sigma^2}(z-\mu)^2\right\} \quad -\infty < z < \infty.$$
- If $Z_1, Z_2, \ldots, Z_n$ are independent with density function $f_1, \ldots, f_n$. Then density of $(Z_1, Z_2, \ldots, Z_n)$ is product of the densities, i.e. $\prod_{i=1}^{n} f_i$
- For $Z_i \sim \mathcal{N}(\mu, \sigma^2)$ and independent. Then we write
$$(Z_1, \ldots, Z_n) \sim \mathcal{N}((\mu, \mu, \ldots, \mu)^\top, \sigma^2 \cdot \mathrm{I})$$

Multivariate normal density

- Identity covariance case

$$f(z_1, z_2, \ldots, z_n) = \frac{1}{\sigma^n (2\pi)^{n/2}} \cdot \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} (z_i - \mu)^2 \right\}$$

Multivariate normal density

- If $Z \sim N(\mu, \Sigma)$, then the density function is

$$f(z) = \frac{1}{\sqrt{|\Sigma|} \cdot (2\pi)^{n/2}} \cdot e^{-(z-\mu)^\top \Sigma^{-1}(z-\mu)}$$

- Density of ϵ?
- What is the density of y, $\hat{e}$, and $\hat{\beta}_{OLS}$?

Properties of Multivariate normal density

- $Z \sim \mathcal{N}(\mu, \Sigma)$, then

$$AZ + c \sim \mathcal{N}(A\mu + c, A\Sigma A^\top)$$

- Density of $\hat{\beta}_{OLS}$?
- Density of $\hat{\beta}_1$?
- Density of $\hat{\beta}_1 - \hat{\beta}_2$?
- Density of $\hat{\beta}_1 - \hat{\beta}_2 + 1$?

- Variance of $\hat{\beta}_1$?

Sampling distribution of $\hat{\boldsymbol{\beta}}$ and $\hat{\epsilon}$ under normality

Assume $\boldsymbol{X}$ is deterministic, and $\mathrm{rank}(\boldsymbol{X}) = p + 1$. $\boldsymbol{H} = \boldsymbol{X}(\boldsymbol{X}'\boldsymbol{X})^{-1}\boldsymbol{X}'$.
Further, assume normality: i.e. $\epsilon_1, \ldots, \epsilon_n$ are iid $N(0, \sigma^2)$.
In particular, the least squares estimate and the residual vectors

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}), \quad \text{and} \quad \hat{\boldsymbol{\epsilon}} \sim N(\boldsymbol{0}, \sigma^2(\boldsymbol{I} - \boldsymbol{H})).$$

Sampling distribution of $\hat{\boldsymbol{\beta}}$ and $\hat{\epsilon}$ under normality

- Under the normality assumption, the least squares estimate

$$\hat{\boldsymbol{\beta}} \sim N(\boldsymbol{\beta}, \sigma^2(\boldsymbol{X}'\boldsymbol{X})^{-1}).$$

- $\hat{\beta}_1 \sim \mathcal{N}(\beta_1, \sigma^2 \cdot [(X^\top X)^{-1}]_{22})$
- Suppose σ^2 is known, then what is the distribution of

$$\frac{\hat{\beta}_1 - \beta_1}{\sigma \cdot \sqrt{[(X^\top X)^{-1}]_{22}}} \sim ?$$

- Why is this useful?
- What happens if we do not know σ^2?

Quadratic forms

- Recall

$$\hat{\sigma}^2 = \frac{1}{n-p-1}\sum_{i=1}^{n}\hat{e}_i^2 = \frac{1}{n-p-1}\hat{e}^\top\hat{e}$$

  Here $\hat{e}$ is the vector of residuals.

- From HW1 $\hat{e} = (I-H)(Y - X\beta^\star) = (I-H)\epsilon$

$$\hat{e}^\top\hat{e} = \epsilon^\top(I-H)\epsilon$$

- Define quadratic forms.

Fisher Cochran's Theorem

Suppose $\boldsymbol{y} \sim N(0, \sigma^2 \boldsymbol{I})$, and P is an idempotent matrix, then we have

$$\frac{1}{\sigma^2} \boldsymbol{y}' P \boldsymbol{y} \sim \chi^2_r,$$

with $r = \mathrm{tr}(P)$.

- Density of χ^2_r

$$f(z) = \frac{1}{2^r \Gamma(r/2)} \cdot z^{r/2-1} e^{-z/2} \quad \text{for } z > 0.$$

- Define the Gamma function. Examples of χ^2-random variables.

Student's t distribution

- Density of student-t random variable with degrees of freedom.

$$f_r(z) = \frac{\Gamma\left(\frac{r+1}{2}\right)}{\sqrt{r\pi} \cdot \Gamma\left(\frac{r}{2}\right)} \cdot \left(1 + \frac{z^2}{r}\right)^{-\frac{r+1}{2}}$$

- Always remeber

$$\frac{\mathcal{N}(0, 1)}{\sqrt{\chi^2(r)/r}} \sim t_r.$$

Here $\mathcal{N}(0, 1)$ and $\chi^2(r)$ are independnt.

Back to unknown σ^2 case

- Recall
$$\frac{\hat{\beta}_1 - \beta_1}{\sigma \cdot \sqrt{[(X^\top X)^{-1}]_{22}}} \sim \mathcal{N}(0, 1)$$
.
- Recall
$$\frac{(n-p-1)\hat{\sigma}^2}{\sigma^2} \sim \chi^2_{n-p-1}$$
- Conclude that
$$\frac{\hat{\beta}_1 - \beta_1}{\hat{\sigma} \cdot \sqrt{[(X^\top X)^{-1}]_{22}}} = \frac{\mathcal{N}(0,1)}{\sqrt{\chi^2(n-p-1)/(n-p-1)}} \sim t_{n-p-1}$$