Lecture Notes: Chapter 5

20/21

5 Estimation

We observe i.i.d. RVs $X_1, X_2, \ldots$ and want to estimate

μ and σ^2.

Unbiased estimators are

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

$$S^2 = \frac{1}{n-1} \sum_{i=1}^{n} (X_i - \overline{X})^2.$$

How good are estimates obtained with them?

5.1. Interval Estimates

As estimates we provide intervals (a, b) such that

$$\mu \in (a,b) \quad \text{or} \quad \sigma^2 \in (a,b)$$

with a "high degree of confidence".

Assumption: $X_1, X_2, \ldots$ are $N(\mu, \sigma^2)$ i.i.d.

(we can sometimes drop this assumption due to the CLT)

Not with "high probability" because we don't know the distribution of μ or σ^2.

Cases:

- σ^2 is known, μ is unknown
- both σ^2 and μ are unknown, interested in μ
- σ^2 is unknown (μ is w/o interest)

Interval Types:

- two-sided intervals: (a, b)
- lower intervals: $(-\infty, b)$
- upper intervals (a, ∞)

5.2 Estimation of two-sided Intervals

Suppose we want to be "95% confident" that $\mu \in (a, b)$.

How should we choose a, b?

Use Case 1: How good is, for given values $x_1, \dots, x_n$ for $X_1, \dots, X_n$, the average $\bar{X}$ as an estimate of the mean $E[X]$?

Idea: Design functions

$$a(x), \quad b(x)$$

with corresponding RVs $A = a(\bar{X})$, $B = b(\bar{X})$ sth

$$P[A \leq \mu \leq B] = 95\%$$

Terminology: $5\% = 100\% - 95\%$ is the confidence level.

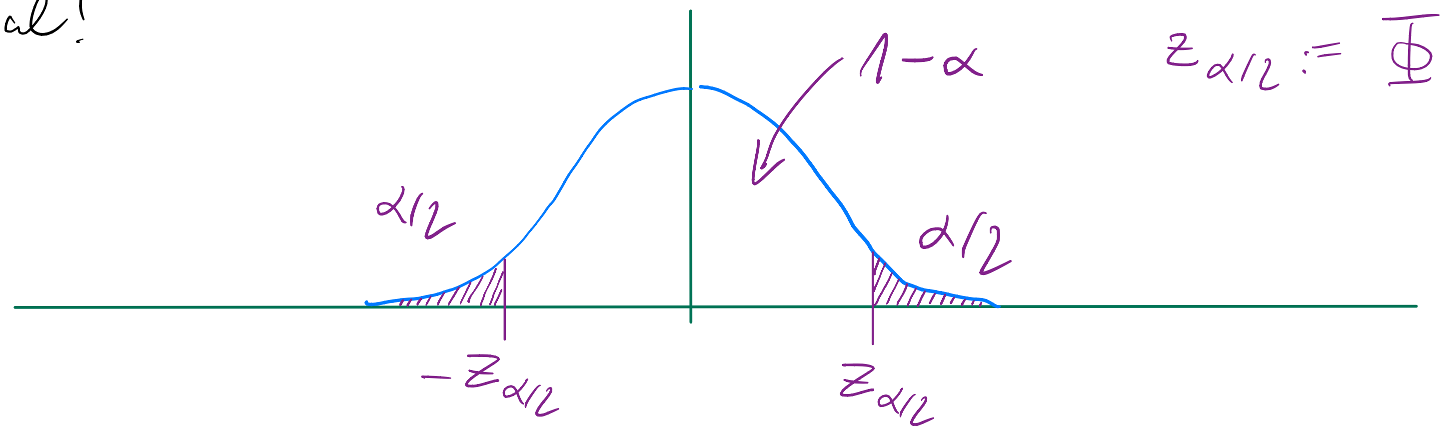
How Choose a, b ?

We consider the problem to determine constants a, b sth

$$P[a \le Z \le b] = 1-\alpha, \qquad Z \sim \mathcal{N}(0,1)$$

Trivial!

$$z_{\alpha/2} := \Phi^{-1}(1 - \alpha/2)$$



Choose $a = -z_{\alpha/2}$, $b = z_{\alpha/2}$

In general define:

$$z_\alpha := \Phi^{-1}(1-\alpha).$$

Next suppose, $X \sim \mathcal{N}(\mu, 1)$, but μ unknown

$\Rightarrow X - \mu \in \mathcal{N}(0,1) \Rightarrow \mu - X \in \mathcal{N}(0,1)$

$$1-\alpha = P\left[-z_{\alpha/2} \le \mu - X \le z_{\alpha/2}\right]$$

$$= P\Big[\underbrace{X - z_{\alpha/2}}_{a(X)} \le \mu \le \underbrace{X + z_{\alpha/2}}_{b(X)}\Big]$$

Note: The interval boundaries are RV, not the mean μ.

Finally, $X_i \sim \mathcal{N}(\mu, \sigma^2) \Rightarrow \bar{X} \sim \mathcal{N}(\mu, \frac{1}{n}\sigma^2)$

$\Rightarrow \bar{X} - \mu \sim \mathcal{N}(0, \frac{1}{n}\sigma^2) \Rightarrow \mu - \bar{X} \sim \mathcal{N}(0, \frac{1}{n}\sigma^2)$

$$\Rightarrow \sqrt{n}\,\frac{\mu - \bar{X}}{\sigma} \sim \mathcal{N}(0,1)$$

$$\Rightarrow 1 - \alpha = P\left[-z_{\alpha/2} \le \sqrt{n}\,\frac{\mu - \bar{X}}{\sigma} \le z_{\alpha/2}\right]$$

$$\Rightarrow \quad 1-\alpha = P\left[-z_{\alpha/2} \le \sqrt{n}\,\frac{\mu-\bar{X}}{\sigma} \le z_{\alpha/2} \right]$$

$$= P\left[-\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \le \mu - \bar{X} \le \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right]$$

$$= P\left[\bar{X} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \le \mu \le \bar{X} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} \right]$$

Proposition: Let $\bar{X} \sim N(\mu, \frac{1}{n}\sigma^2)$. Let

- $a(x) = x - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$
- $b(x) = x + \frac{\sigma}{\sqrt{n}} z_{\alpha/2}$.

Then

$$P\left[a(\bar{X}) \le \mu \le b(\bar{X}) \right] = 1-\alpha$$

We say that $(a(\bar{x}), b(\bar{x}))$ is a $(1-\alpha)\cdot 100\%$ confidence interval for the mean, for a given $\bar{x}$.

5.3 Estimation of One-sided intervals

Proposition: Let $\bar{X} \sim N(\mu, \frac{1}{n}\sigma^2)$ and let

- $a(x) = x - \frac{\sigma}{\sqrt{n}} z_\alpha$
- $b(x) = x + \frac{\sigma}{\sqrt{n}} z_\alpha$

Then

- $P[\, a(\bar{X}) \leq \mu \,] = 1 - \alpha$
- $P[\, \mu \leq b(\bar{X}) \,] = 1 - \alpha$

$(a(\bar{x}), \infty)$ and $(-\infty, b(\bar{x}))$ are upper and lower confidence intervals

5.4. Fixing the interval length

Suppose we want a confidence interval $(a(\bar{x}), b(\bar{x}))$

- for $(1-\alpha)\cdot 100\,\%$
- of length at most L,

How many observations are needed?

- The length of $(a(\bar{x}), b(\bar{x}))$ is

$$b(\bar{x}) - a(\bar{x}) = \bar{x} + \frac{\sigma}{\sqrt{n}} z_{\alpha/2} - \left(\bar{x} - \frac{\sigma}{\sqrt{n}} z_{\alpha/2}\right)$$

$$= 2\frac{\sigma}{\sqrt{n}} z_{\alpha/2}$$

- Hence $b(\bar{x}) - a(\bar{x}) \leq L \iff 2\frac{\sigma}{\sqrt{n}} z_{\alpha/2} \leq L$

$$\iff \sqrt{n} \geq 2\frac{\sigma}{L} z_{\alpha/2} \iff n \geq \left(2\frac{\sigma}{L} z_{\alpha/2}\right)^2$$

5.5 Where we need the Normal

We only used the assumption that

$$\bar{X} \sim \mathcal{N}(\mu, \tfrac{1}{n}\sigma^2).$$

Due to the CLT, this is also approximately true for averages of non-normal RVs.

Rule of Thumb: For $n \geq 30$, we can assume a normal distribution of $\bar{X}$

5.6 Confidence Interval for μ with unknown σ^2

Reduction to $\sqrt{n}\,\frac{\bar{X}-\mu}{\sigma} \sim \mathcal{N}(0,1)$ does not work.

However, if $X_1, X_2, \ldots$ are i.i.d. $\mathcal{N}(\mu, \sigma^2)$, then

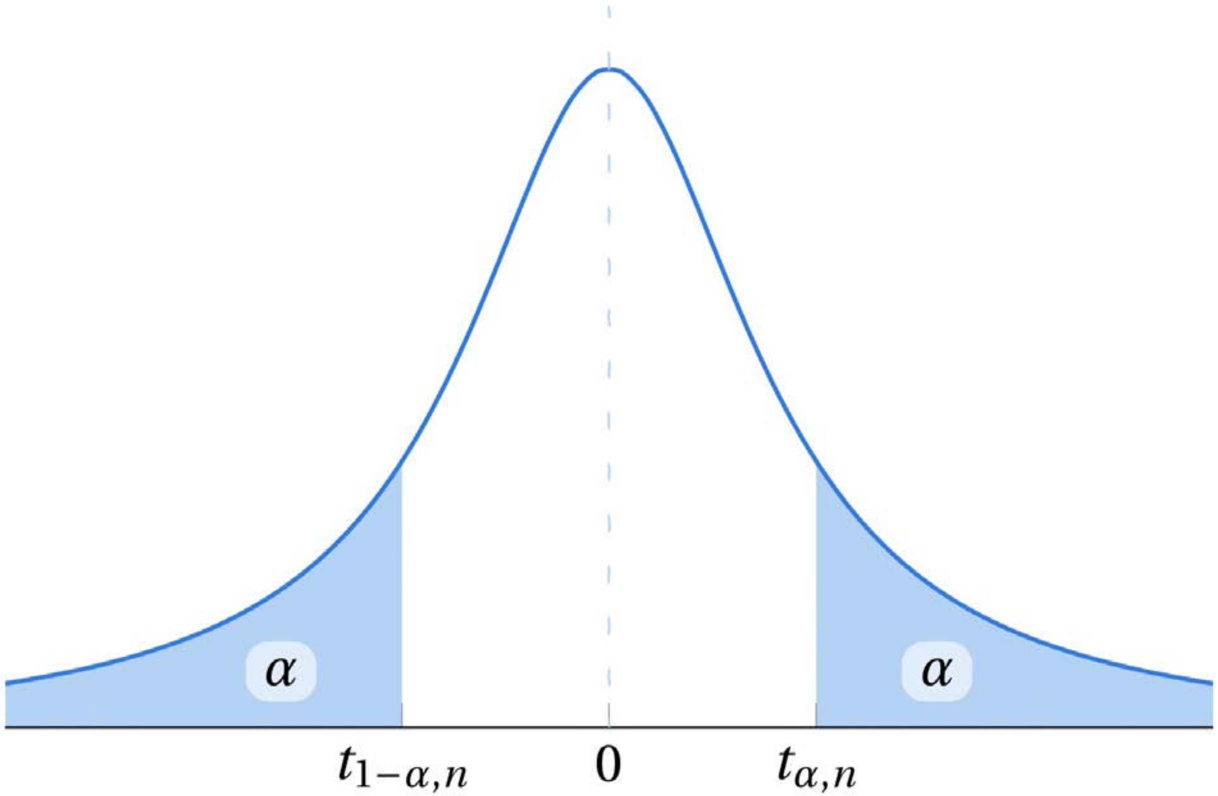
$$\sqrt{n}\,\frac{\bar{X}-\mu}{S} \sim t_{n-1}$$

Analogously to z_α, where

$$P[Z \le z_\alpha] = 1-\alpha$$

define $t_{\alpha,n}$ by

$$P[t_n \le t_{\alpha,n}] = 1-\alpha$$



The areas before $-t_{\alpha,n}$ and after $t_{\alpha,n}$ have size α.

Analogous theory holds with

- $a(x) = x - \frac{s}{\sqrt{n}} t_{\alpha/2, n-1}$
- $b(x) = x + \frac{s}{\sqrt{n}} t_{\alpha/2, n-1}$

$\Bigg\}$ for 2-sided intervals $(a(\bar{x}), b(\bar{x}))$

- $a(x) = x - \frac{s}{\sqrt{n}} t_{\alpha, n-1}$ for $(a(\bar{x}), \infty)$

- $b(x) = x + \frac{s}{\sqrt{n}} t_{\alpha, n-1}$ for $(-\infty, b(\bar{x}))$

Remarks: (1) This rests on the assumption that the X_i are normal.

(2) Usage of t_{n-1} only necessary for small n ($n \lessapprox 30$), since

$$t_n \longrightarrow N(0,1)$$

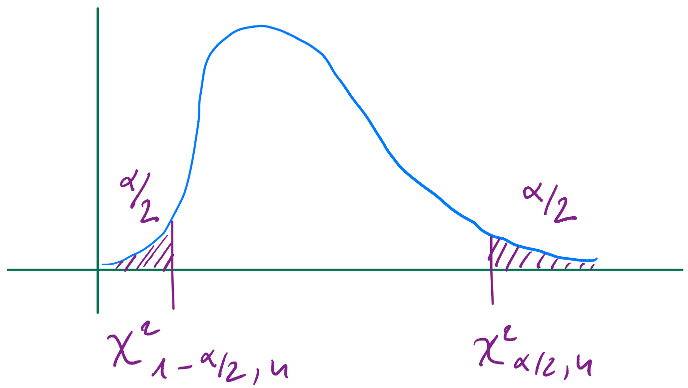
(3) Whether t_{n-1} is needed or $N(0,1)$ will do, depends on the confidence level:

- the higher the confidence, the smaller α, the further $t_{\alpha,n}$ to the tail
- the slower convergence to the normal

5.7 Confidence Intervals for σ^2

Recall:

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



By definition,

$$1-\alpha = P\left[\chi^2_{1-\alpha/2,n-1} \le (n-1)\frac{S^2}{\sigma^2} \le \chi^2_{\alpha/2,n-1}\right]$$

$$= P\left[\frac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}} \le \sigma^2 \le \frac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}\right]$$

The 2-sided boundary functions are then

- $a(S^2) = \dfrac{(n-1)S^2}{\chi^2_{\alpha/2,n-1}}$
- $b(S^2) = \dfrac{(n-1)S^2}{\chi^2_{1-\alpha/2,n-1}}$

One-sided boundary functions are

- $a(S^2) = \dfrac{(n-1)\,S^2}{\chi^2_{\alpha,n-1}}$

- $b(S^2) = \dfrac{(n-1)\,S^2}{\chi^2_{1-\alpha,n-1}}$