Dec 6^{th}

- Today: last class
- Next Friday 4-6:45pm take-home (online) final exam
- Course Experience Survey

Maximum Likelihood is one method to get a point estimate (MLE)

Another method to get a point estimate is Method of Moments (MoM)

What are moments?

$E[X] \rightarrow 1^{st}$ moment mean

$E[X^2] \rightarrow 2^{nd}$ moment variance

$E[X^3] \rightarrow 3^{rd}$ moment skewness

$\vdots$

$E[X^k] \rightarrow k^{th}$ moment

$E[X]$ is approximated by $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

More generally: estimate the k^{th} moment $E[X^k]$, we can take a random sample of size n and $E[X^K] \approx \frac{1}{n}\sum_{i=1}^{\infty} X_i^K$

* many downsides to MoM

Moment estimator: $\hat{\theta}_{MOM}$ are obtained by using the moments

ex Let $X_1, \dots, X_n \overset{i.i.d.}{\sim} \exp(\lambda)$

collect n samples, all i.i.d. w/ exp R.V. where the λ is unknown. $(\theta = \lambda)$

Note for exp R.V. mean $= E[X] = \frac{1}{\lambda}$ (prior ch)

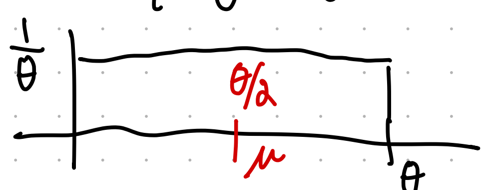
But we approximate $E[X]$ with $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$

$$\Rightarrow E[X] \approx \bar{X} = \frac{1}{\lambda}$$

$$\hat{\lambda} = \frac{1}{\bar{X}} = \frac{1}{\frac{1}{n}\sum_{i=1}^{n} X_i}$$

---

ex $X_1, \dots, X_n \sim \text{Unif}(0, \theta) = \begin{cases} \frac{1}{\theta} & 0 \le x \le \theta \\ 0 & \text{otherwise} \end{cases}$



estimate θ using MoM:

$$E[X] = \frac{\theta}{2} \approx \bar{X}$$

$$\Rightarrow \hat{\theta} = 2\bar{X} = 2 \frac{1}{n} \sum_{i=1}^{n} X_i$$

↳ unreasonable since $\hat{\theta} > \max(X_1, \dots, X_n)$ while each $X_i < \theta$ ($2\bar{X}$ might not be bigger than all X_i's)

Let $X_1, \dots, X_{10} \sim$ i.i.d w/ distribution below

$$f(x;\theta) = \begin{cases} (\theta+1)x^{\theta} & 0 \le x \le 1 \\ 0 & \text{otherwise} \end{cases}$$

where $-1 < \theta$.

$X_1 = .92$ $X_2 = .79$ $X_3 = .90$ $X_4 = .65$ $X_5 = .86$
$X_6 = .47$ $X_7 = .73$ $X_8 = .97$ $X_9 = .94$ $X_{10} = .77$

$$\bar{X} = .8$$

Use MLE and MoM to estimate θ.

MoM: $E[X] = \int_0^1 x(\theta+1)x^{\theta}\,dx = \frac{(\theta+1)x^{\theta+2}}{\theta+2}\Big|_{0}^{x=1}$

$$= \frac{\theta+1}{\theta+2} \approx \overline{X}$$

$$\frac{\hat{\theta}+1}{\hat{\theta}+2} \approx 0.8$$

$$\hat{\theta}+1 = \frac{4}{5}(\hat{\theta}+2)$$

$$\hat{\theta} = 3$$

MLE: $f(x_1, \dots, x_{10}; \theta) = \underbrace{(\theta+1)^n (x_1 x_2 \cdots x_{10})^\theta}_{\text{Likelihood}}$

ln trick:

$$\ln\left[(\theta+1)^n (x_1 x_2 \cdots x_{10})^\theta\right]$$

$$n \ln(\theta+1) + \ln\left(x_1^\theta x_2^\theta \cdots x_{10}^\theta\right)$$

$$\mathcal{L}(\theta) = n \ln(\theta+1) + \theta \sum_{i=1}^{10} \ln(x_i)$$

$$\frac{d\mathcal{L}}{d\theta} = \frac{n}{\theta+1} + \sum_{i=1}^{10} \ln(x_i) = 0$$

$$\underline{\hat{\theta}_{MLE}} = \frac{-n}{\sum \ln(x_i)} - 1 \approx \frac{10}{-(\ln(x_1) + \dots + \ln(x_{10}))} - 1$$

$$= 3.12$$

end point estimates

Previously $\hat{\theta}$, a point estimate for θ, was one single value. However, $\hat{\theta}$ alone provides limited info b/c we cannot determine how close it is to the true value θ.

Instead, now: use intervals

For example, instead of reporting $\hat{\theta} = 3.12$

we can provide an interval

$$[2.9, 3.2] \qquad [\hat{\theta}_l, \hat{\theta}_h]$$

that contains (we hope) θ.

We want say how "confident" we are w/ this interval that it contains θ, true value

ex Collect data $X_1, \dots, X_n \overset{i.i.d}{\sim} N(\mu, \sigma^2)$

μ → unknown

σ^2 → known

maybe not realistic but good starting point

What values of μ are believeable given data $X_1, \dots, X_n$?

↳ want an interval

MoM estimator: $\bar{X} = \frac{1}{n} \sum_{i=n}^{n} X_i$

Standardize:

know, unknown

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

σ ← know

$\sqrt{n}$ ← know

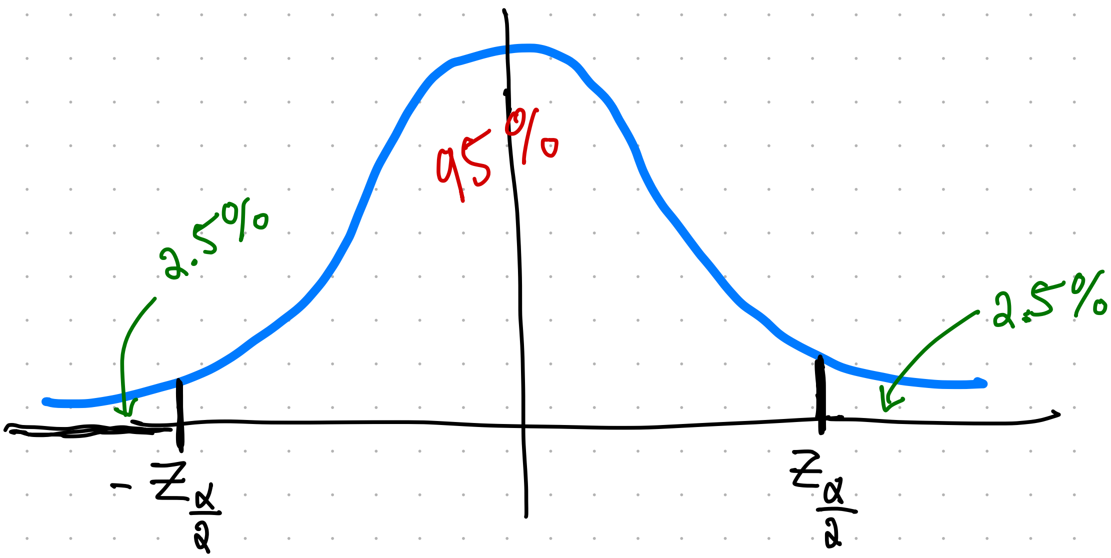
$$\frac{\sum X_i - n\mu}{\sigma\sqrt{n}} \quad \text{CLT}$$

divide by n:

$$\frac{\frac{1}{n}\sum_i X_i - \mu}{\frac{\sigma\sqrt{n}}{n}} \qquad \frac{\sigma\sqrt{n}}{n} = \frac{\sigma}{\sqrt{n}}$$

We want a 95% confidence interval

Notation: $\alpha = 0.05$ want $(1-\alpha)100\%$ C.I.



$$P\left(-1.96 \le \frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le 1.96\right) = 0.95$$

long: $P\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le 1.96\right) - \underbrace{P\left(\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \le -1.96\right)} = 0.95$

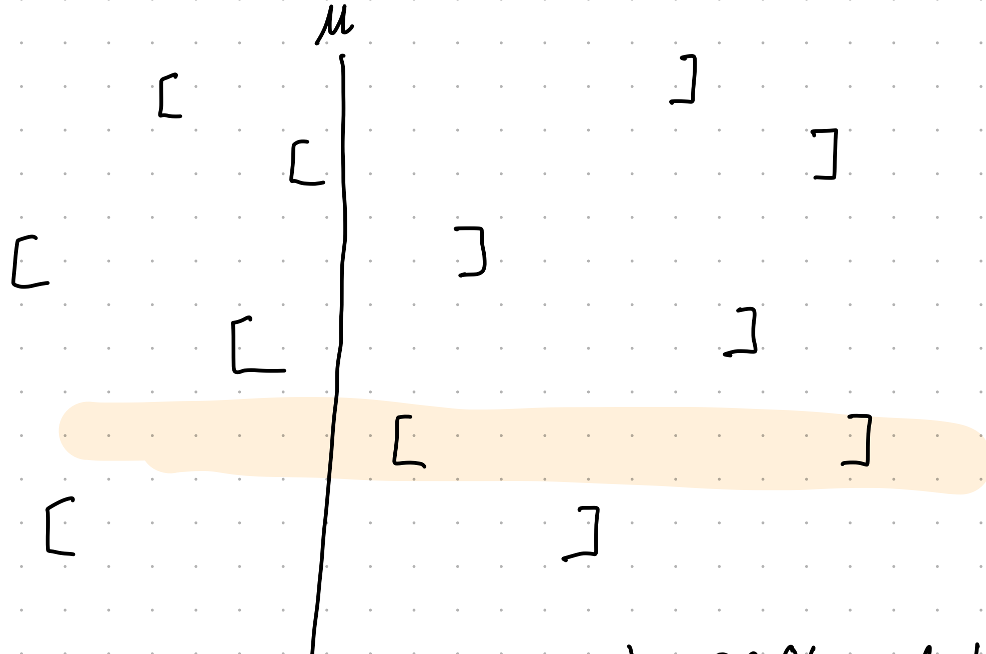
$$P\left(-1.96\frac{\sigma}{\sqrt{n}} \le \bar{X}-\mu \le 1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

know #s → constant

$$P\left(\bar{X}-1.96\frac{\sigma}{\sqrt{n}} \le \mu \le \bar{X}+1.96\frac{\sigma}{\sqrt{n}}\right) = 0.95$$

random interval

The prob that $\left[\bar{X} - \frac{1.96\sigma}{\sqrt{n}}, \bar{X} + \frac{1.96\sigma}{\sqrt{n}}\right]$ interval will cover μ is 0.95. (95%)



Run experiment numerous times & 95% of time, the true unknown μ will be in the interval.

↳ 5% of the intervals will not have μ

↳ you won't know when

Summary: A $100(1-\alpha)\%$ Confidence Interval (C.I) for μ when σ^2 is known:

$$\left[\bar{X} - Z_{\frac{\alpha}{2}} \cdot \frac{\sigma}{\sqrt{n}}, \bar{X} + Z_{\frac{\alpha}{2}} \frac{\sigma}{\sqrt{n}}\right]$$

Common Z_α, $Z_{\alpha/2}$

$100(1-\alpha)\%$	α	$\alpha/2$	$Z_{\alpha/2}$
80%	.2	.1	1.282
90%	.1	.05	1.645
95%	.05	.025	1.96
99%	.01	.005	2.576

What about if σ is NOT known. Assume n is large (CLT applies)

Before:

$$P\left(\bar{X} - Z_{\alpha/2}\frac{\sigma}{\sqrt{n}} \le \overset{\mu}{\theta} \le \bar{X} + Z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right) = 1-\alpha$$

if σ is unknown:

1. create an upper bound for σ
2. estimate σ

Ex (Opinion Polling) We would like to estimate the portion of people who plan to vote for candidate A in election. A voter is chosen at random uniformly among voters and ask who they will vote for.

if yes to Cand A $\longrightarrow X=1$
if no to Cand A $\longrightarrow X=0$

$$f_X(x) = \begin{cases} 1 & \text{yes A} \\ 0 & \text{no A} \end{cases}$$

$$X \sim \text{Bernoulli}(\theta)$$

Then we want a confidence interval for this structure.

$E[X_i] = \theta$ — we want to estimate the mean of distribution

Note: $Var(X_i) = \sigma^2 = \theta(1-\theta)$
$p(1-p)$ for Bernoulli

Thus, to find σ, we need to know θ. But θ is the parameter we want to estimate in the first place. Try to bound σ.

We want to bound $f(\theta) = \theta(1-\theta)$ (find max)

$$f(\theta) = \theta - \theta^2$$

$$f'(\theta) = 1 - 2\theta = 0$$

$\theta = \frac{1}{2}$ max occurs

$$\Rightarrow \sigma^2 = f(\theta) = \theta(1-\theta) \le \underbrace{f\left(\tfrac{1}{2}\right)}_{\text{max}} = \tfrac{1}{2}\left(1-\tfrac{1}{2}\right) = \tfrac{1}{2}\cdot\tfrac{1}{2} = \tfrac{1}{4}$$

$$\Rightarrow \sigma = \sqrt{\tfrac{1}{4}} = \tfrac{1}{2} \qquad \sigma_{max} = \tfrac{1}{2}$$

C.I: $\left[\bar{X} - Z_{\alpha/2}\frac{\sigma_{max}}{\sqrt{n}}, \bar{X} + Z_{\alpha/2}\frac{\sigma_{max}}{\sqrt{n}}\right]$

$$\Rightarrow \left[\bar{X} - Z_{\alpha/2}\frac{1}{2\sqrt{n}}, \bar{X} + Z_{\alpha/2}\frac{1}{2\sqrt{n}}\right]$$ is a $100(1-\alpha)\%$ C.I for θ

Bernoulli in general

when μ, σ unknown

ex Polling: How large does n need to be so that we obtain a 90% (95%) C.I. with 3% margin of error?

$$P(\bar{X} - 0.03 \leq \theta \leq \bar{X} + 0.03) \geq .90$$

Modelled by Bernoulli: (use C.I. derived!)

$$\left[\bar{X} - \frac{Z_{\alpha/2}}{2\sqrt{n}}, \bar{X} + \frac{Z_{\alpha/2}}{2\sqrt{n}}\right]$$ valid $100(1-\alpha)\%$ C.I for θ

Match: $\frac{Z_{\alpha/2}}{2\sqrt{n}} = 0.03$ $\qquad \alpha = .1 \longrightarrow Z_{\frac{\alpha}{2}} = 1.645$ chart

$$\frac{1.645}{2\sqrt{n}} = 0.03$$ Solve for n

$$\left(\frac{1.645}{2 \cdot 0.03}\right)^2 = n$$

$$n \geq 752$$

For 95% C.I: $\alpha = .05 \longrightarrow Z_{\alpha/2} = 1.96$
chart

$$\frac{Z_{\alpha/2}}{2\sqrt{n}} = 0.03$$

$$\left(\frac{1.96}{2 \cdot 0.03}\right)^2 = n$$

$$n \geq 1068$$

Sometimes we want to test the hypothesis...

ex drug company wants to if a drug is effective

A statistical hypothesis is an assertion or claim about one or more population characteristics

ex proportion of students majoring in Science $\leq .6$
height of students ≥ 170 cm
height of students $\sim N(170, 5)$

Kinds of Hypothesis

Null Hypothesis: H_0 initially assumed to be true

Alternate Hypothesis : H_1 hypothesis compared against H_0

ex $X \sim N(\mu, 36)$ and μ is either 50 or 55

$H_0 : \mu = 50$

$H_1 : \mu = 55$

Decide whether to reject H_0 using a rule:

ex: C & C' partition (disjoint)

$C = \{(X_1, \dots, X_n) : \bar{X} \geq 53\} \rightarrow$ reject H_0 in favor of H_1

data

$C' = \{\underbrace{(X_1, \dots, X_n)}_{\text{subject to randomness}} : \bar{X} < 53\} \rightarrow$ do not reject H_0

subject to randomness

error

decision

truth (unknown)		reject H_0	do not reject H_0
	H_0 correct	type I error	✓
	H_1 correct	✓	type II error

Notation:

$P(\text{type I error}) = \text{significance level} = \alpha$

$P(\text{type II error}) = \beta$

$X \sim N(\mu, 36)$

ex
$H_0: \mu = 50$
$H_1: \mu = 55$

reject if $\bar{X} \geq 53$
do not reject if $\bar{X} < 53$

take a sample of $n = 16$

$\alpha = P(\text{type I error}) = P(\text{reject } H_0 \text{ when } H_0 \text{ is true})$

$= P(\text{reject } H_0 \mid \mu = 50)$

$\alpha = P(\bar{X} \geq 53 \mid \mu = 50)$

Standardize

$\sigma^2 = 36 \quad n = 16$
$\sigma = 6 \quad \sqrt{n} = 4$

$$\alpha = P\left(\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \geq \frac{53 - \mu}{\frac{\sigma}{\sqrt{n}}} \,\middle|\, \mu = 50\right)$$

$$\alpha = P\left(\frac{\bar{X} - 50}{\frac{6}{4}} \geq \frac{53 - 50}{\frac{6}{4}}\right)$$

CLT

$$\approx P(\underset{\text{standard normal}}{Z} \geq 2) = 1 - P(Z \leq 2)$$

$$= 1 - .9772 = 0.0228$$

This means if in reality the null hyp is true, $\mu = 50$, and if we take a sample size of 16 then 2.28% we will mistakenly reject the null hyp (type I error).

Then $\beta = P(\text{type II error}) = P(\text{do not reject } H_0 \mid \underbrace{H_0 \text{ incorrect}}_{\text{alternative } H_1 \text{ is correct}})$

$$= P(\bar{X} < 53 \mid \mu = 55)$$

$$= P\left(\frac{\bar{X} - 55}{6/4} < \frac{53-55}{6/4}\right) \approx P\left(Z < -\frac{4}{3}\right)$$

$$= 0.0913$$

In reality, if null is incorrect (so $\mu = 55$) and we take a sample size of $n = 16$ & calculate the sample mean and based on that, you reject/no do not reject H_0, then 9.13% you will not reject the null (incorrect)