MESSAGE FOR MY 2023/24 STUDENTS:

This slide pack was designed to teach students about pivotal functions. It was written before I took over EC400 (where I introduced pivotal functions to you). So there is nothing new in here. I'm just making it available in case anyone wanted a refresher. All the best. RAGUIR (24/11/23)

EC402 (20/21) - Common Question - Pivots / C.I.s

Ragvir's Speaking Notes (Unofficial Content)

ANOTHER SLIDE PACK?

"RAGVIR, YOU'RE THE BEST. I WISH YOU WERE MY UNDERGRAD PROF."

Aww... shucks... thanks! Flattery helps. Here is some free info. on intervals.

"RAGVIR, WHY SHOULD MY REAL UG STATS TEACHER BE IN PRISON?"

Because he/she taught you C.I.s without mentioning pivots, and it is (or should be) a criminal offence.

"RAGVIR, BUT THAT'S JUST ANOTHER EXTRA CONCEPT YOU ARE THROWING AT ME"

Well, the alternative is to just memorise things like

"if σ^2 is known, it's a z-score but
if σ^2 is unknown, it becomes a t-stat. Yay!",

and then pass your UG exams but as soon as you hit questions like Q1 of PS7, you don't know where to even begin. Up to you.

"RAGVIR... YOU'RE ACTUALLY THE WORST."

Sigh!... I know ... I know...

WHAT ARE WE TRYING TO DO?

POINT ESTIMATION

- let $x_i \overset{iid}{\sim} N(\mu, \sigma^2)$ for $i = 1, \dots, N$. We saw (in "Pushing Limits") that $\bar{x}_N := \frac{1}{N}\sum_{i=1}^{N} x_i$ is the ML estimator for μ.
- This gives us a point estimate. What if we want an interval estimate (eg: $\bar{x}_N \pm$ something)?

INTERVAL ESTIMATION

- A $100(1-\alpha)\%$ C.I. for μ is given by $[T_L(x), T_u(x)]$ where $T_L(x) < \bar{x}_N$ and $T_u(x) > \bar{x}_N$ are chosen such that

$$P\left(T_L(x) < \mu < T_u(x)\right) = 1-\alpha$$

for some given α.

make sure you can translate into plain english. It is **NOT** "the probability that μ is between $T_L(x)$ and $T_u(x)$", please!

WHY DO MY EC402 STUDENTS STRUGGLE WITH THIS?

So how do we come up with $T_L(x)$ and $T_u(x)$? Well, let's look at this letter from your UG teacher...

Dear ex-Student of mine, 17/Nov/2020

Hope you are well. I'm in prison right now for crimes against statistics. But let me remind you what I taught you in the good old days...

- If X_is are Normal, I said look at "$\sqrt{N}(\bar{X}_N - \mu)/\sigma$", but I never told you 'why'.
- If σ^2 is known, I said use a "Z score" but I never told you why. Just use the $N(0,1)$ tables, and be satisfied.
- If σ^2 is unknown, we can use a "t-statistic" but I never told you why. Just remember to replace σ with s, ok? Oh...and use the t tables.
- When the situation is more complex (eg. difference of group means, etc.), then

I gave you additional expressions to memorise but I don't remember those myself so how can I ask you to remember?!

To SUMMARISE, if $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ for $i = 1, \dots, N$, then...

- If you know σ^2, get a $100(1-\alpha)\%$ C.I. as

$$\bar{X}_N \pm \frac{\sigma}{\sqrt{N}} z_{1-\frac{\alpha}{2}}$$

, using $N(0,1)$ tables; AND

- If you don't know σ^2, get a $100(1-\alpha)\%$ C.I. as

$$\bar{X}_N \pm \frac{s}{\sqrt{N}} t_{N-1, 1-\frac{\alpha}{2}}$$

, using t_{N-1} tables.

NOTE: If you get any other question that I did not cover whatever I have taught you will be impossible for you to apply since I taught it so badly.

Hope it helps! Or Not!

Mr. Generic UG Teacher

PIVOTAL FUNCTIONS: SOME ST202 NOTES BELOW...

- for a proper definition, look up any stats book. Something like this...

Definition of a pivotal function: Consider a sample Y with density $f_Y(y|\theta)$ and suppose that we are interested in constructing an interval estimator for θ. A function $G = G(Y, \theta)$ of Y and θ is a **pivotal function** for θ if its distribution is known and does not depend on θ.

↳ My ST202 students typically didn't find this enough for intuition so I made the notes below for them.

Hope it helps you too. After reading it, see my answer to Q1, PS7. Now, can you see how I came up with various pivots?

PIVOTAL FUNCTIONS - A FRIENDLY INTRODUCTION

03/03/19

☺

-RAGVIR

QUESTION : HOW DO I THINK ABOUT PIVOTS? I GET THE THEORY, BUT HOW TO APPROACH THE HOMEWORK?

just an example to fix ideas. ← SETUP : $X_i \overset{iid}{\sim} N(\mu, \sigma^2)$ with σ^2 known for $i=1,\dots,N$.

GOAL : TO CONSTRUCT CI'S FOR μ, UNKNOWN

(I) Q: So why is a Pivot useful? Why is it even called a Pivot?

A: Recall how we construct CI's...

eg: Given the "setup" above, we can say (usual ST102 stats tables stuff):
Since we have normality,
we can always find a q_1 and q_2 s.t.

$$P\left(q_1 < \frac{\sqrt{N}(\bar{x}-\mu)}{\sigma} < q_2\right) = 1-\alpha$$

where α is typically some small number like 0.05 or 0.01.

Indeed from stats tables, $q_{1,\alpha/2=0.025} = -1.96$

$q_{2,1-\frac{\alpha}{2}=0.975} = 1.96$

So, we know that

$$P\left(-1.96 < \frac{\sqrt{N}(\bar{x}-\mu)}{\sigma} < 1.96\right) = 0.95$$

Now, think of the highlighted part as the Pivot. Why? Because we will isolate μ in the middle and move all the other "observed" stuff to the other sides of the inequalities. We are "pivoting" around this object, right? (I will wave my hands a lot and show what I mean in the seminar!)

i.e. $$P\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{N}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{N}}\right) = 0.95$$

- Make sure you can work this out very clearly yourself.
- I will show steps in the seminar.

So, we use the above knowledge as motivation to come up with an interval estimator for μ.

i.e. a $100(1-\alpha)\%$ C.I for μ is given by:

$$\bar{x} \pm 1.96\,\sigma/\sqrt{N}$$

(II) Q: The above was an illustrative example. Let's talk about how to find appropriate pivots from a practical standpoint.

NOTE: Matteo has given you the OFFICIAL definition. You need to know this. What follows below is just a loose intuitive "Ragvir's guidelines" type of discussion for you.

A: Keep the "GOAL" in mind. Now think about guessing a suitable pivot.

(a) - It needs to contain μ, right? Otherwise, we would do tonnes of algebra and there would be no probability statement in the end involving our unknown parameter, the very object of our interest!

(b) - Its distribution should be KNOWN and should not depend on any unknown parameter. Otherwise, how would we look up a lovely stats table and dig out values of the relevant quantiles q_1 and q_2?!

(c) - It should not contain anything unknown in it. Otherwise, after you do the algebra, you'd get say:

$$P\left(\bar{x} - 1.96\frac{\sigma}{\sqrt{N}} < \mu < \bar{x} + 1.96\frac{\sigma}{\sqrt{N}}\right) = 0.95$$

... then you'd try to plug the lower and upper limits of the interval into your calculator. Here, I assumed σ was known (see first slide) so I had no problem. But say σ was unknown, you'd be utterly stuck!!

[Test yourself: What would be an appropriate pivot if σ^2 was really not known in my example?]

So (a), (b), and (c) allow you to construct a suitable pivot. There is an additional point which allows you to construct the "best" pivot.

(d) - It should not "waste" info. If you have info, use it! For example, if you have N observations, use them all. If you know σ^2 don't try to estimate it. This will give you the "tightest" interval estimators (for a given α).

I really hope this helps. I will explain all this in my seminars anyway.

See you soon! -RAGVIR.

BACK TO EC402 PS7, Q1

- I wanted a P.I. for y_{12}:

  ⇒ I needed a pivot which (i) I knew the distribution of (fully);

  (ii) contained y_{12};

  (iii) contained nothing else unknown;

  (iv) didn't waste any info.

  Indeed, $(\hat{y}_{12} - y_{12}) / \hat{var}^{1/2}(\hat{y}_{12} - y_{12} | X)$ worked perfectly.

- I wanted a pivot for M_{12}:

  The above would be useless. Indeed,

  $(\hat{y}_{12} - M_{12}) / \hat{var}^{1/2}(\hat{y}_{12} - M_{12} | X)$ was suitable.

[Same logic works also for Q2]