Victoria Lombardi

Math 369 9/21/20

Lecture 7

DGP: $X_{1}, \ldots, X_{n_1} \overset{iid}{\sim} N(\theta_1, \sigma_1^2)$ indep of $X_{21}, \ldots X_{2n_2} \overset{iid}{\sim} N(\theta_2, \sigma_2^2)$

Now we don't assume we know σ_1^2 & σ_2^2 and we use the sample variances to estimate them

$$S_1^2 := \frac{1}{n_1 - 1} \sum_{i=1}^{n_1} (X_{1i} - \bar{X}_1)^2, \quad S_2^2 := \frac{1}{n_2-1} \sum_{i=1}^{n_2} (X_{2i} - \bar{X}_2)^2$$

Under $H_0: \theta_1 - \theta_2 = 0$

$$\Rightarrow \frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim T_{df} \, ?$$ But no....

This was pointed by Behrens (1929) & Fisher (1935) because they discovered this distribution, it's called the Behrens-Fisher distribution (and this is called the Behrens-Fisher problem).

$$\frac{\hat{\theta}_1 - \hat{\theta}_2}{\sqrt{S_1^2/n_1 + S_2^2/n_2}} \sim \text{BehrensFisher}(\ldots)$$

They tried to work out its PDF but they couldn't and at some point they gave up & conjectured that it was impossible. In 1966, it was proven that it has a closed form solution. And, it was published in 2018.

In 1946/7, Welch & Satterthwaite found a T approximation which is very good & still used today (p 314 C&B):

$$df = \frac{\left(\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}\right)^2}{\frac{S_1^4}{n_1^2(n_1-1)} + \frac{S_2^4}{n_2^2(n_2-1)}}$$

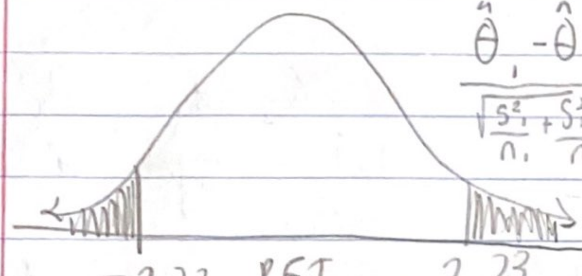
Using this T_{df} is known as Welch's t-test or "unequal variances t test".

②

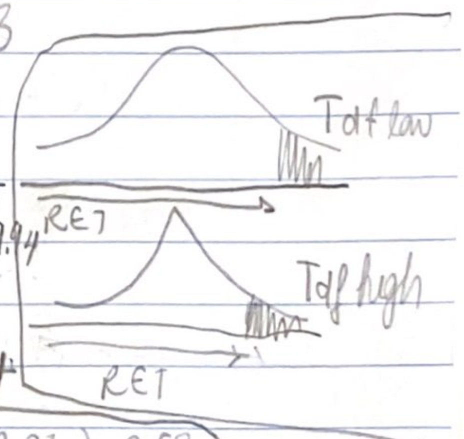
$n_1 = 10$, $\bar{X}_1 = 70.5$, $S_1 = 2.07$ } male
$n_2 = 6$, $\bar{X}_2 = 62.3$, $S_2 = 2.25$ } female

$$\underbrace{\frac{\hat\theta_1 - \hat\theta_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}}_{} \sim T_{df} \nearrow df = \frac{1.27^2}{\frac{2.07^4}{10^2(9)} + \frac{2.25^4}{6^2(5)}} = \frac{1.62}{0.163} = 9.94$$

$$SE = \sqrt{\frac{2.07^2}{10} + \frac{2.25^2}{6}} = \sqrt{1.27} = 1.13$$



$$\frac{\hat\theta_1 - \hat\theta_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \Big| H_0 \sim T_{9.94}$$



$F_{T_{9.94}}(-2.23) = 2.5\%$

$$\frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{2.07^2}{10} + \frac{2.25^2}{6}}} = \frac{8.2}{1.13} = 7.27$$

Obs, std, statistic

$\Rightarrow$ Reject!

MIDTERM 1 ↑ Midterm 2 ↓

$X_1, \dots, X_n \overset{iid}{\sim} DGP(\theta_1, \theta_2, \dots \theta_K)$ K is # parameters

We've previously seen estimators $\hat\theta = w(X_1, \dots, X_n)$ e.g. $\hat\theta = \bar{X}$, $\hat\sigma^2 = \frac{1}{n}\Sigma(X_i - \bar{X})^2 \dots$

How did we get the function w? Where did it come from? There are many strategies to create estimators, the first we'll study is called "Methods of Moments" (MM) and it was used by Karl Pearson in the late 1890's. We know the DGP so we know which θ's we want to estimate. We now need an algorithm to generate w.

Def. The K^{th} moment of a rv is $E[X^K]$.
The first moment is $\mu_1 := E[X^1]$, the second is $\mu_2 := E[X^2]$, etc...

We define the "sample moments" as: $\hat{\mu}_K := \frac{1}{n}\sum_{i=1}^{n} X_i^K$

The first sample moment is the "sample average" sample mean.

$\hat{\mu}_1 = \frac{1}{n}\sum X_i = \bar{X}$

Pearson's idea is to "match moments to parameters" If...

$\mu_1 = \alpha_1(\theta_1, \ldots, \theta_K),\ \mu_2 = \alpha_2(\theta_1, \ldots, \theta_K), \ldots\ldots$

$\mu_K = \alpha_K(\theta_1, \ldots, \theta_K)$

and

$\theta_1 = \gamma_1(\mu_1, \ldots, \mu_K),\ \theta_2 = \gamma_2(\mu_1, \ldots, \mu_K) \ldots$

$\theta_K = \gamma_K(\mu_1, \ldots, \mu_K)$ system of equations

$\Rightarrow \hat{\theta}_i^{mm} = \gamma_i(\hat{\mu}_1, \ldots \hat{\mu}_K)$

MM pretty much always gives you an estimator, but it is rarely a "great" estimator & sometimes produces totally wrong answers

$X_1, \ldots, X_n \overset{iid}{\sim} N(\overset{\mu}{\theta_1}, \overset{\sigma^2}{\theta_2})$ We want the MM estimator for both θ_1 (mean) & θ_2 (variance) in the iid normal DGP

true for ALL DGP's $\theta_1 = E[X] = \gamma_1(\mu_1, \mu_2) = \mu_1 \Rightarrow \hat{\theta}_1^{mm} = \hat{\mu}_1 = \bar{X}$

variance(X): $\theta_2 = \gamma_2(\mu_1, \mu_2) = \mu_2 - \mu_1^2 \Rightarrow \hat{\theta}_2^{mm} = \hat{\mu}_2 - \hat{\mu}_1^2 = \frac{1}{n}\sum X_i^2 - \bar{X}^2 = \hat{\sigma}^2$

$\hat{\sigma}^2 = \frac{1}{n}\sum(X_i - \bar{X})^2 = \frac{1}{n}\sum(X_i^2 - 2X_i\bar{X} + \bar{X}^2) = \frac{1}{n}\sum X_i^2 - \frac{1}{n}2\bar{X}(n\bar{X}) + \frac{1}{n}n\bar{X}^2$

$= \frac{1}{n}\sum X_i^2 - \bar{X}^2$

$X_1, \ldots, X_n \overset{iid}{\sim} Bin(\theta_1^{=n}, \theta_2^{=p})$ both θ_1, θ_2 unknown

We want to estimate both θ_1 (which is commonly denoted n) and θ_2 (which is commonly denoted p). Ecologists love this estimation problem b/c it's part of the "capture-recapture" problem to estimate population size of wildlife. Each data point is the result of catching a certain # of fish in a time interval (e.g. 1 hr of fishing). Once you catch a fish you release it & re-cast. Everytime a fish encounters the hook it's a Bern(θ_2) that it bites and you

CAtch #.

θ_2 is the propensity to bite and θ_1 is the # of individual fish-hook encounters in the time-period (e.g. 1hr). (4)

lets develop MM estimators for both θ_1 + θ_2

$E[X] = \mu_1 = \alpha_1(\theta_1, \theta_2) = \theta_1\theta_2 \Rightarrow \theta_1 = \frac{\mu_1}{\theta_2}$ ← solve for gamma$_1$ + gamma$_2$

$\mu_2 = Var[X] + \mu_1^2 = \theta_1\theta_2(1-\theta_2) + \theta_1^2\theta_2^2 = \alpha_2(\theta_1, \theta_2)$

$= \theta_1\theta_2 - \theta_1\theta_2^2 + \theta_1^2\theta_2^2$

$= \frac{\mu_1}{\theta_2}\theta_2 - \frac{\mu_1}{\theta_2}\theta_2^2 + \frac{\mu_1^2}{\theta_2^2}\theta_2^2$

$= \mu_1 - \mu_1\theta_2 + \mu_1^2 = \mu_2$

$\Rightarrow \mu_2 - \mu_1^2 - \mu_1 = \mu_1\theta_2 \Rightarrow \theta_2 = \frac{\mu_1^2 + \mu_1 - \mu_2}{\mu_2} = \frac{\mu_1 - (\mu_2 - \mu_1^2)}{\mu_1}$

$\theta_1 = \frac{\mu_1}{\frac{\mu_1 - (\mu_2 - \mu_1^2)}{\mu_1}} = \frac{\mu_1^2}{\mu_1 - (\mu_2 - \mu_1^2)}$,

$\Rightarrow \hat{\theta}_1^{mm} = \frac{\hat{\mu}_1^2}{\hat{\mu}_1 - (\hat{\mu}_2 - \hat{\mu}_1^2)}$, $\hat{\theta}_2^{mm} = \frac{\hat{\mu}_1 - (\hat{\mu}_2 - \hat{\mu}_1^2)}{\hat{\mu}_1}$,

$\hat{\theta}_1^{mm} = \frac{\bar{X}^2}{\bar{X} - \hat{\sigma}^2}$ $\hat{\theta}_2^{mm} = \frac{\bar{X} - \hat{\sigma}^2}{\bar{X}}$

$n=5$, $\bar{x} = <3, 7, 5, 5, 6> \Rightarrow \bar{x} = 5.2$, $\hat{\sigma}^2 = 2.64$

$\hat{\theta}_1^{mm} = \frac{5.2^2}{5.2 - 2.64} = 10.56$, $\hat{\theta}_2^{mm} = \frac{5.2 - 2.64}{5.2} = .49$

$n=5$, $\bar{x} = <3, 7, 5, 11, 6> \Rightarrow \bar{x} = 6.4$, $\hat{\sigma}^2 = 10.56$

$\hat{\theta}_1^{mm} = \frac{6.4^2}{6.4 - 10.56} = -9.8$, $\hat{\theta}_2^{mm} = \frac{6.4 - 10.56}{6.4} = -.65$

Obviously, n can't be negative and p must be a probability so these estimates are nonsensical. MM estimators are sometimes really bad... but they make for a nice place to start...