Lecture 20

Topics:

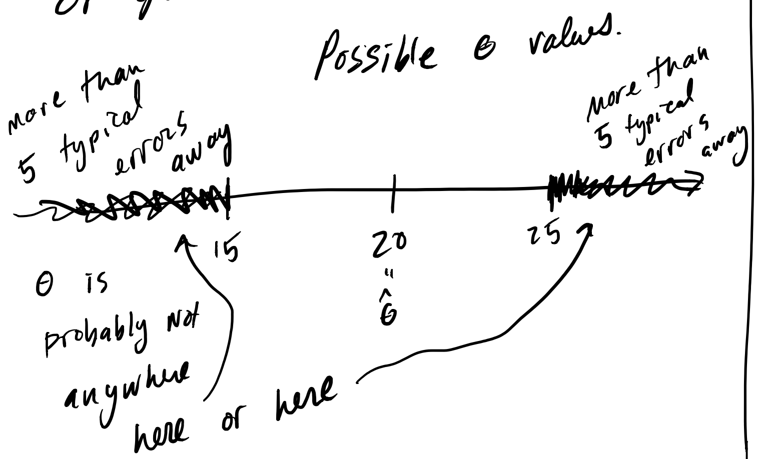
I) Hypothesis testing.

Hypothesis testing is essentially a formal way of ruling out possible values of a population parameter θ when observing an estimate $\hat{\theta}$.

e.g. You want to estimate some unknown θ. You collect data and get an estimate $\hat{\theta} = 20.0$

Suppose the $MSE(\hat{\theta}) = 1.0$

Then almost certainty the true θ isnt less than 15 or greater than 25



Note: the logic is a bit sublte

We know $\hat{\theta}$ is a random draw from a distribution that is close to θ (by about ± 1)

$\uparrow$ $\sqrt{MSE}$

e.g. suppose $\hat{\theta} \sim N(\theta, 2)$ where θ is unknown.

You sampled one draw from $N(\theta, 2)$ & got

$$\hat{\theta} = -105.7$$

Can you rule out some values for θ?

Z-scores help you determine what values of θ are plausable:

$$\text{z-score} = \frac{\hat{\theta} - E(\hat{\theta})}{sd(\hat{\theta})} = \frac{-105.7 - \theta}{\sqrt{2}}$$

Remember:

- z-score of ± 2 is Normal
- z-score of ± 3 is kinda unlikely
- z-score of ± 4 is rare
- z-score of ± 5 is extremely rare.

hypothetical values of θ	z-score of $\hat{\theta} = -105.7$	Plausable value of $\hat{\theta}$?
$\theta = 1.2$	$\frac{-105.7 - 1.2}{\sqrt{2}} = -75$	No way
$\theta = -95.2$	$\frac{-105.7 - (-95.2)}{\sqrt{2}} = -7.49$	No way
$\theta = -110$	$\frac{-105.7 - (-110)}{\sqrt{2}} = 3.04$	Somewhat unlikely
$\theta = -108$	$\frac{-105.7 - (-108)}{\sqrt{2}} = 1.62$	plausable.

Another way to phrase it: ③

* $\theta = -95.2$ is inconsistent with the data $\hat{\theta} = -105.7$
* $\theta = -108$ is consistent with the data $\hat{\theta} = -105.7$

To use z-scores for testing plausability (or consistency) of a particular θ, you need to compute $E(\hat{\theta})$ & $sd(\hat{\theta})$ as a function of θ.

$E(\hat{\theta})$ & $sd(\hat{\theta})$ depend on the unknown θ.

to test a particular value of θ, just plug it in where ever it occurs in $E(\hat{\theta})$ & $sd(\hat{\theta})$.

e.g. Huge list of 0's & 1's:

$$x_1, x_2, \ldots, x_{100000000}.$$

You want to estimate the proportion of 1's in the list (call this number θ). So you sample (w/rep) 2000 members from the list & get

$$X_1, X_2, \ldots, X_{2000}.$$

where

$$\hat{\theta} = \frac{X_1 + \cdots + X_{2000}}{2000} = \bar{X} = 0.67.$$

Is it plausable that $\theta = 0.5$? ④

$$\begin{pmatrix} \text{z-score for} \\ \hat{\theta} = 0.67 \text{ when} \\ \theta = 0.5 \end{pmatrix} = \frac{0.67 - E(\hat{\theta})}{sd(\hat{\theta})}$$

$$= \frac{0.67 - \theta}{\sqrt{\frac{\theta(1-\theta)}{2000}}}$$

replace θ with the possible value your testing, 0.5 in this case

$$= \frac{0.67 - 0.5}{\sqrt{\frac{0.5(1-0.5)}{2000}}}$$

$$= 15.2$$

$\therefore$ $\theta = 0.5$ would be totally inconsistent with our observation that $\hat{\theta} = 0.67$

What if $E(\hat{\theta})$ or $sd(\hat{\theta})$ depend on other unknown population parameters?

In this case it often works to "plug in" an estimate of the other unknowns when computing the z-score.

e.g. Another huge list of numbers

$$y_1, y_2, \ldots, y_N \quad \text{where } N = 5\times 10^{22}$$

You want to estimate

$$\mu = \frac{1}{N}\sum_{i=1}^{N} y_i = \bar{y}$$

i.e. the list average.

Define the list s.d. as

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(y_i-\bar{y})^2}.$$

Sometimes called the population s.d.

To construct $\hat{\mu}$ you randomly sample (w/Rep) $n = 25$ numbers from the list & get $Y_1, Y_2, \ldots, Y_n$ with

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} Y_i = 18.72.$$

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(Y_i-\bar{Y})^2} = 3.5$$

sometimes called the sample s.d.

(one reason why we divide by $n-1$ for $\hat{\sigma}$ is that it makes $\hat{\sigma}^2$ an unbiased est of σ... this type of adjustment is more important when doing "Regression". We do it here so you start getting used to it).

Question: is $\mu = 17.0$ a plausable value for μ?

Ans:

The z-score for $\hat{\mu} = 18.72$ when $\mu = 17.0$ is given by

$$z = \frac{\hat{\mu} - E(\hat{\mu})}{sd(\hat{\mu})}$$

$$= \frac{18.72 - \mu}{\sqrt{\frac{\sigma^2}{n}}}$$

replace this with the value we are testing.

$$= \frac{18.72 - 17}{\left(\frac{\sigma}{\sqrt{25}}\right)}$$

Now plug in $\hat{\sigma} = 3.5$ for σ

$$\cong \frac{18.72 - 17}{\left(\frac{3.5}{\sqrt{25}}\right)}$$

$$= 2.06$$

$\therefore$ The z-score of our observation ($\hat{\mu} = 18.72$, $\hat{\sigma} = 3.5$) when $\mu = 17.0$ is totally plausable.

i.e. we can't rule out that the true value of μ is 17.0

e.g. Two huge lists of numbers ⑦

$x_1, x_2, \dots,$

$y_1, y_2, \dots$

Let μ_x = the Ave of the x_i's

μ_y = the Ave of the y_i's.

σ_x = the "population s.d." of the x_i's

σ_y = the "population s.d." of the y_i's.

You want to estimate

$$\mu_x - \mu_y = ?$$

So you sample 25 numbers from the first list & get

$$X_1, X_2, \dots, X_{25}$$

also sample 20 numbers from the second list & get

$$Y_1, Y_2, \dots, Y_{20}.$$

Suppose this data results in

$$\hat{\mu}_x = \bar{X} = -2.5$$

$$\hat{\mu}_y = \bar{Y} = 1.23$$

$$\left.\begin{aligned}\hat{\sigma}_x &= 10.2\\ \hat{\sigma}_y &= 5.1\end{aligned}\right\}$$ sample standard deviations of the two samples.

⑧

Question: is it plausable that $\mu_x - \mu_y = 0$, i.e. that $\mu_x = \mu_y$?

Answer:

Lets compute the z-score of the estimate $\hat{\mu}_x - \hat{\mu}_y = -2.5 - 1.23$ under the hypothetical that $\mu_x - \mu_y = 0$.

Note: MF1 & MF2 give

$$E(\hat{\mu}_x - \hat{\mu}_y) = \mu_x - \mu_y$$

$$sd(\hat{\mu}_x - \hat{\mu}_y) = \sqrt{\frac{\sigma_x^2}{25} + \frac{\sigma_y^2}{20}}$$

$$\therefore \quad z\text{-score} = \frac{(\hat{\mu}_x - \hat{\mu}_y) - (\mu_x - \mu_y)}{\sqrt{\frac{\sigma_x^2}{25} + \frac{\sigma_y^2}{20}}}$$

hypothetical value of $\mu_x - \mu_y$

$$= \frac{(-2.5 - 1.23) - (0)}{\sqrt{\frac{\sigma_x^2}{25} + \frac{\sigma_y^2}{20}}}$$

$$\approx \frac{(-2.5 - 1.23) - (0)}{\sqrt{\frac{10.2^2}{25} + \frac{5.1^2}{20}}}$$

plug in $\hat{\sigma}_x = 10.2$ & $\hat{\sigma}_y = 5.1$ for σ_x & σ_y

$$= -1.59$$

$\therefore$ $\mu_x - \mu_y = 0$ is totally consistent with the data.