14.320: Recitation 2

Martina Uccioli

February 16, 2023

Note: see handwritten notes (at the end of this handout) for solutions

Contents

1 Logs

Why is $\frac{A-B}{B} \approx \log(A) - \log(B)$?

1. Taylor expansion of $\log(x)$ around 1
   - $y = \log(x)$
   - $y = \log(x)$ and $y = x - 1$
2. Replace x with $\frac{A}{B}$

```
.
. ttest loguhe, by(immig)

Two-sample t test with equal variances
------------------------------------------------------------------------------
   Group |     Obs        Mean    Std. err.   Std. dev.   [95% conf. interval]
---------+--------------------------------------------------------------------
       0 |  52,300    3.300842    .0031697     .724887    3.294629    3.307055
       1 |  14,879    3.137078    .0064108    .7819809    3.124512    3.149644
---------+--------------------------------------------------------------------
Combined |  67,179    3.264571     .002859    .7410341    3.258967    3.270175
---------+--------------------------------------------------------------------
    diff |            .1637641    .0068562                .1503259    .1772022
------------------------------------------------------------------------------
    diff = mean(0) - mean(1)                                      t =  23.8855
H0: diff = 0                                     Degrees of freedom =    67177

    Ha: diff < 0                 Ha: diff != 0                 Ha: diff > 0
 Pr(T < t) = 1.0000         Pr(|T| > |t|) = 0.0000          Pr(T > t) = 0.0000
```

Figure 1: From Lecture Notes 3

calculator

2 Chebyshev's Inequality

For any random variable, X_i, and any positive constant, c:

$$P(|X_i - \mu_x| \geq c\sigma_x) \leq \frac{1}{c^2}$$

can also be written as

$$P(|X_i - \mu_x| \geq c) \leq \frac{\sigma_x^2}{c^2}$$

3 Law of Large Numbers

Let $\bar{X}_n$ be the sample mean in a random sample of size n. The LLN says

$$\lim_{n\to\infty} P\{|\bar{X}_n - \mu_x| \geq \epsilon\} = 0$$

4 Montecarlo Simulations and Confidence Intervals

See do file

A Expectations of a function (Law of the Unconscious Statistician)

See handwritten notes

① LOGS

want to show: $\frac{A-B}{B} \approx \log(A) - \log(B)$

(1) Taylor expansion of $\log(x)$ around $x=1$

$$\log(x) \approx \underbrace{\log(1)}_{0} + \underbrace{\frac{\partial \log(x)}{\partial x}\Big|_{x=1}}_{= 1/x |_{x=1} = 1} (x-1) + \dots$$

$$\log(x) \approx x-1$$

(2) $x = A/B$

$$\log\left(\frac{A}{B}\right) \approx \frac{A}{B} - 1$$

$$\boxed{\log(A) - \log(B) \approx \frac{A-B}{B}}$$

example from class:

$$\log(A) - \log(B) = .1637\ldots$$

$$\log\left(\frac{A}{B}\right) = .1637\ldots$$

$$\left(\frac{A}{B}\right) = e^{.1637\ldots}$$

$$\frac{A-B}{B} = e^{.1637\ldots} - 1 \Rightarrow \frac{A-B}{B} = .177\ldots$$

$\approx$

② Chebyshev's Inequality

- first we prove Markov's inequality, but just as a learning tool (you don't need to learn this) —

MARKOV'S INEQUALITY:

X is nonnegative random variable, then $P(X \geq \alpha) \leq \frac{E[X]}{\alpha}$

$$E[X] \underset{\text{by LIE}}{=} E[X \mid X \geq \alpha] P(X \geq \alpha) + \underbrace{E[X \mid X < \alpha]}_{\geq 0 \text{ ← } X \text{ is} \geq 0} \underbrace{P(X < \alpha)}_{\geq 0 \text{ ← probabilities always} \geq 0}$$

$$\geq \underbrace{E[X \mid X \geq \alpha]}_{\geq \alpha \text{ ← we are conditioning on } X \geq \alpha} P(X \geq \alpha)$$

$$\boxed{E[X] \geq \alpha P(X \geq \alpha)}$$

Chebyshev's

for <u>any</u> r.v X. and any $c > 0$

$$P(|X_i - \mu_x| \geq c\sigma_x) \leq \frac{1}{c^2}$$

$$\sigma_x^2 = E[(x_i - \mu_x)^2]$$

by LIE

$$= E[(x_i-\mu_x)^2 \mid |X_i - \mu_x| \geq c\sigma_x] P(|X_i - \mu_x| \geq c\sigma_x)$$

$$+ \underbrace{E[(x_i-\mu_x)^2 \mid |X_i - \mu_x| < c\sigma_x]}_{\geq 0} \underbrace{P(|X_i - \mu_x| < c\sigma_x)}_{\geq 0}$$

$$\geq \underbrace{E[(x_i-\mu_x)^2 \mid \underbrace{|X_i - \mu_x| \geq c\sigma_x}_{\Rightarrow (x_i-\mu_x)^2 \geq c^2\sigma_x^2}]}_{\geq c^2\sigma_x^2} P(|X_i - \mu_x| \geq c\sigma_x)$$

$$\boxed{\sigma_x^2 \geq c^2\sigma_x^2 P(|X_i - \mu_x| \geq c\sigma_x)}$$

$$\frac{\sigma_x^2}{c^2\sigma_x^2} \qquad \frac{}{c\sigma_x^2}$$

$$\frac{1}{c^2} \geq P(|X_i - \mu_x| \geq c\sigma_x)$$

③ Law of large Numbers

(i) recall chebyshev's, for a random variable Y_i

$$P(|Y_i - \mu_y| \geq c) \leq \frac{\sigma_{y_i}^2}{c^2}$$

(2) Apply chebyshev's to the random variable $\bar{X}_n$

$Y_i = \bar{X}_n$

$$E[\bar{X}_n] = E\left[\frac{1}{n}\sum X_i\right] = \frac{1}{n}\sum E[X_i]$$

$$= \frac{1}{n} n \cdot \mu_x = \mu_x$$

$$var(\bar{X}_n) = var\left(\frac{1}{n}\sum X_i\right) = \frac{1}{n^2}\sum var(X_i)$$

$$= \frac{1}{n^2} n \sigma_x^2 = \frac{\sigma_x^2}{n}$$

$$\boxed{P(|\bar{X}_n - \mu_x| \geq \epsilon) \leq \underbrace{\frac{\sigma_x^2}{n}\frac{1}{\epsilon^2}}}$$

$n \to \infty$ $\to 0$

(A) Appendix A: Law of the Unconscious Statistician

want to show : $E[g(x)] = \int_{-\infty}^{\infty} g(x) f_x(x) dx$

[if you simply apply the definition, you get
$E[g(x)] = \int_{-\infty}^{\infty} g(x) f_{g(x)}(g(x))\, dg(x)$]

(1) start from a random variable y:

$E[y] = \int_{-\infty}^{\infty} y f_y(y)\, dy$

(2) change of variable $y = g(x)$ $\quad x = g^{-1}(y)$

$\boxed{dy = g'(x)\, dx}$

$dy = g'(g^{-1}(y))\, dx$

$dx = \frac{1}{g'(g^{-1}(y))} dy$

$F_Y(y) = P(Y \leq y)$

$F_Y(g(x)) = P(g(X) \leq y)$

$= P(X \leq g^{-1}(y))$

$= F_X(g^{-1}(y))$

$f_y(y) = \frac{dF(y)}{dy} = \frac{dF_X(g^{-1}(y))}{dg^{-1}(y)} \underbrace{\frac{dg^{-1}(y)}{dy}}_{= \frac{1}{g'(g^{-1}(y))}}$

$$\boxed{f_y(y) = f_x(x) \frac{1}{g'(x)}}$$

(3) plug (2) in (1)

$$E[y] = \int_{-\infty}^{\infty} y f_y(y)\, dy$$

$$E[g(x)] = \int_{-\infty}^{\infty} g(x)\, f_x(x) \frac{1}{\cancel{g'(x)}}\, \cancel{g'(x)}\, dx$$

$$= \int_{-\infty}^{\infty} g(x)\, f_x(x)\, dx$$