② 4(c)

CSE 544, Spring 2017, Probability and Statistics for Data Science

Assignment 4: Statistical Inference Due: 4/05, in class

(7 questions, 70 points total)

I/We understand and agree to the following:

(a) Academic dishonesty will result in an 'F' grade and referral to the Academic Judiciary.

(b) Late submission, beyond the 'due' date/time, will result in a score of 0 on this assignment.

(write down the name of all collaborating students on the line below)

---

1. More on Wald's test **(Total 5 points)**

Suppose the null hypothesis is H_0: $\theta = \theta_0$, but the true value of θ is θ_*. Show that, under Wald's test, the probability of a Type II error is $\varphi(\frac{\theta_0-\theta_*}{\hat{se}} + z_{\alpha/2}) - \varphi(\frac{\theta_0-\theta_*}{\hat{se}} - z_{\alpha/2})$.

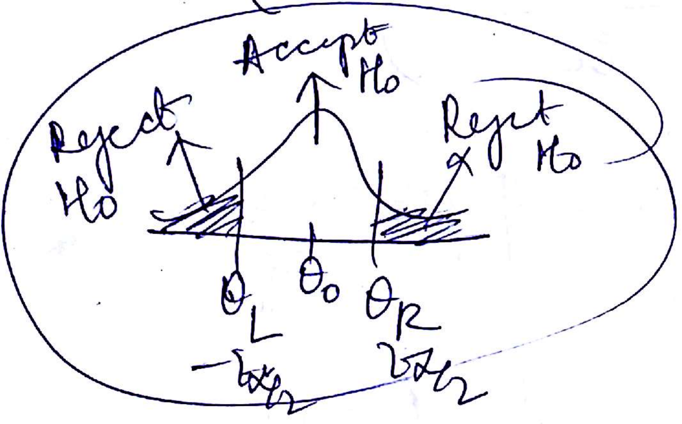
(Hints: (i) might help to draw a figure; (ii) think about the distribution of the estimate.)

P(type II error)

~~= P(reject Ho Even when H_1~~ fail to

~~= P(fail to reject Ho~~

P(donot reject Ho | $\theta = \theta_*$)
When Ho is false



~~θ_R~~

Reject Ho when $|W| > z_{\alpha/2}$

$\Rightarrow W > z_{\alpha/2} \Rightarrow \frac{\theta - \theta_0}{\hat{se}} > z_{\alpha/2}$

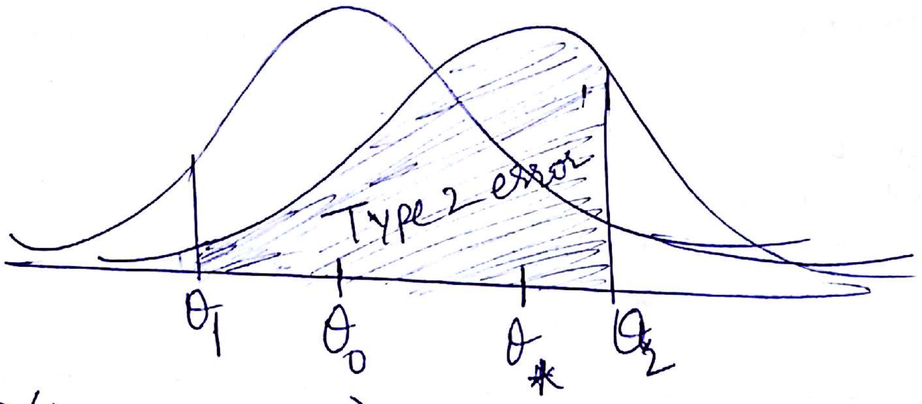
$\Rightarrow \theta > \boxed{\theta_0 + \hat{se}\, z_{\alpha/2} = \theta_R}$

$\Rightarrow W < -z_{\alpha/2} \Rightarrow \frac{\theta - \theta_0}{\hat{se}} < -z_{\alpha/2}$

$\theta < \theta_0 - \hat{se}\, z_{\alpha/2} = \theta_L$

Scanned by CamScanner

true value of $\theta = \theta_*$



$P(\text{type 2 error})$

$= P(\text{Do not reject } H_0 \mid \theta = \theta_*)$

$= P(\theta_L < \theta < \theta_R \mid \theta = \theta_*)$

$$= P\left(\frac{\theta_L - \theta_*}{\hat{se}} < Z < \frac{\theta_R - \theta_*}{\hat{se}}\right)$$

$$= \phi\left(\frac{\theta_R - \theta_*}{\hat{se}}\right) - \phi\left(\frac{\theta_L - \theta_*}{\hat{se}}\right)$$

$$= \phi\left(\frac{\theta_0 + \hat{se}\, z_{\alpha/2} - \theta_*}{\hat{se}}\right) - \phi\left(\frac{\theta_0 - \hat{se}\, z_{\alpha/2} - \theta_*}{\hat{se}}\right)$$

$$= \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} + z_{\alpha/2}\right) - \phi\left(\frac{\theta_0 - \theta_*}{\hat{se}} - z_{\alpha/2}\right)$$

Scanned by CamScanner

2. Posterior for Normal

Let $X_1, X_2, \dots, X_n$ be distributed as Normal(θ, σ²), where σ is assumed to be known. You are also given that the prior for θ is Normal(a, b²). **(Total 10 points)**

(a) Show that the posterior of θ is Normal(x, y²), such that:

$x = \frac{b^2\bar{X}+se^2 a}{b^2+se^2}$ and $y^2 = \frac{b^2 se^2}{b^2+se^2}$; where $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $se^2 = \sigma^2/n$. (7 points)

(Hint: less messier if you ignore the constants, but please justify why you can ignore them)

(b) Compute the (1-α) posterior interval for θ. (3 points)

(a) $f(x_1 x_2 \dots x_n | \theta) \propto \exp\left(-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}\right)$

Prior $g(\theta) \propto \exp\left(-\frac{n(\bar{x}-a)^2}{2b^2}\right)$

Posterior of θ

$h(\theta | x_1 x_2 \dots x_n) = L_n(\theta) \times f(\theta)$

$\propto f(x_1 x_2 \dots x_n | \theta)\, g(\theta)$

$\propto \exp\left(-\frac{n(\bar{x}-\theta)^2}{2\sigma^2}\right) \exp\left(\frac{-(\theta-a)^2}{2b^2}\right)$

$\propto \exp\left(-\frac{n(\bar{x}-\theta)^2}{2\sigma^2} - \frac{(\theta-a)^2}{2b^2}\right)$

$\propto \exp\left(\frac{-nb^2(\bar{x}-\theta)^2 - \sigma^2(\theta-a)^2}{2\sigma^2 b^2}\right)$

$\propto \exp\left\{-\left(\frac{(nb^2+\sigma^2)\theta^2 - 2(n\bar{x}b^2+a\sigma^2)\theta + nb^2\bar{x}^2 + a^2\sigma^2}{2\sigma^2 b^2}\right)\right\}$

$\propto \exp\left\{-\left(\frac{(b^2+se^2)\theta^2 - 2(\bar{x}b^2+a\,se^2)\theta + b^2\bar{x}^2 + a^2 se^2}{2se^2 b^2}\right)\right\}$

Scanned by CamScanner

$$\propto \exp\left\{ - \frac{\left(\theta - \left(\frac{a\,se^2 + \bar{x}\, b^2}{b^2+se^2}\right)\right)^2}{2\left(\frac{b^2 se^2}{b^2+se^2}\right)} \right\} \times \text{constant}$$

$$\sim \text{Normal}\left(\frac{a\,se^2+\bar{x}\,b^2}{b^2+se^2}, \frac{b^2 se^2}{b^2+se^2}\right)$$

$$\sim \text{Normal}(x, y^2)$$

$$\Rightarrow \boxed{x = \frac{a\,se^2+\bar{x}\,b^2}{b^2+se^2}, \quad y = \frac{b^2 se^2}{b^2+se^2}}$$

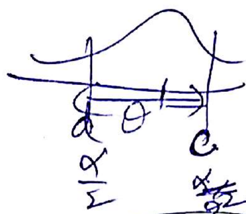
where $se^2 = \frac{\sigma^2}{n}$

$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

(b) Chose c & d s.t.

$P(\theta < c \mid x_1 x_2 \ldots x_n) = \frac{\alpha}{2}$

$P(\theta > d \mid x_1 x_2 \ldots x_n) = \frac{\alpha}{2}$



Mean $= x$
Var ~~dekken~~ $(\sigma^2) = y^2$

$\Rightarrow P(\theta < c \mid x_1 x \ldots x_n)$

$= P\left(\frac{\theta - x}{y} < \frac{c-x}{y}\right) = P\left(Z < \frac{c-x}{y}\right)$

$= \frac{\alpha}{2}$

$\Rightarrow \frac{c-x}{y} = z_{\alpha/2} \Rightarrow \boxed{c = x + z_{\alpha/2}\, y}$

$-z_{\alpha/2} = \frac{d-x}{y} \Rightarrow \boxed{d = x - z_{\alpha/2}\, y}$

$\Rightarrow (1-\alpha)$ posterior interval $= (x \pm z_{\alpha/2}\, y)$.

3. **Conjugate posterior for Poisson** **(Total 5 points)**

Let $X_1, X_2, \dots, X_n$ be distributed as Poisson(λ). Let Gamma(α, β) be the prior of λ, where the pdf of Gamma(α, β) is such that f(x) is proportional to $x^{\alpha-1}e^{-x\beta}$. Show that the posterior is also a Gamma and find its parameters. Feel free to ignore the constants and conclude that the posterior is a Gamma if its form resembles that of f(x) above.

$X_1 X_2 \dots X_n \rightarrow \text{Poisson}(\lambda)$

Gamma (α,β) = prior of λ

pdf of Gamma $(\alpha,\beta) \Rightarrow f(x) \propto x^{\alpha-1} e^{-x\beta}$

posterior of $\lambda \propto$ Liklihood $\times$ prior

$f(\theta|X_1 X_2 \dots X_n) \propto L_n(\theta) f(\theta) \Rightarrow f(\lambda|X_1 X_2 - X_n) \propto L_n(\lambda) f(\lambda)$

$$L_n(\theta) = \prod_{i=1}^{n} \text{Poisson}(\lambda)$$

$$= \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{\lfloor x_i} = \frac{e^{-n\lambda} \lambda^{\sum x_i}}{\prod \lfloor x_i}$$

$$\Rightarrow \quad f(\theta|X_1 X_2 \dots X_n) \propto \frac{e^{-n\lambda} \lambda^{\sum x_i}}{\prod \lfloor x_i} \times \lambda^{\alpha-1} e^{-\lambda\beta}$$

$$\propto \frac{1}{\prod \lfloor x_i} e^{-n\lambda-\lambda\beta} \lambda^{\sum x_i + \alpha - 1}$$

$$\sim \text{Gamma}\left(\sum x_i + \alpha,\ n+\beta\right)$$

4(c)

4. Practice with MLE **(Total 10 points)**

(a) Let $X_1, X_2, \ldots, X_n$ be distributed as Poisson(λ). Find the MLE of λ. (3 points)

(b) Let $X_1, X_2, \ldots, X_n$ be distributed as Binomial(n, p). Find the MLE of p, assuming n is fixed. (4 points)

(c) Let $X_1, X_2, \ldots, X_n \sim$ Normal(θ, 1). Let $\delta = E[I_{X_1>0}]$. Use the Equivariance property to show that the MLE of δ is $\varphi(\frac{1}{n}\sum_{i=1}^{n} X_i)$. You can assume the MLE of the Normal as derived in class. (3 points)

(a) Poisson(λ)

$$P(X=x) = \frac{e^{-\lambda}\lambda^{x}}{\lfloor x}$$

$$L(\lambda) = \prod_{i=1}^{n} \frac{e^{-\lambda}\lambda^{x_i}}{\lfloor x_i} = \frac{\lambda^{\sum x_i} e^{-n\lambda}}{\prod \lfloor x_i}$$

$$\log L(\lambda) = \sum x_i \log\lambda - n\lambda - \log(\prod \lfloor x_i)$$

max by finding derivative

$$\frac{d}{d\lambda}\log(L(\lambda)) = \frac{\sum x_i}{\lambda} - n = 0$$

$$\boxed{\hat{\lambda} = \frac{\sum x_i}{n} = \bar{x}}$$

(b) Bin(n,p)

$$P(X=x) = {}^{n}C_{x}\, p^{x}(1-p)^{n-x}$$

for $X_1 X_2 X_3 - X_n$ iid Bin random var will have joint frequency function that is product of marginal frequency functions.

$\Rightarrow$ liklihood, $L(p) = \prod_{i=1}^{n} {}^{n}C_{x_i}\, p^{x_i}(1-p)^{n-x_i}$

$$= \left(\prod {}^{n}C_{x_i}\right) p^{\sum x_i}(1-p)^{(n^2-\sum x_i)}$$

log liklihood, $\log L(p) = \log \prod {}^{n}C_{x_i} + \sum x_i \log p + (n^2 - \sum x_i)\log(1-p)$

find max by findng derivative,

$$\frac{d}{dp} L'(p) = 0 + \frac{\sum x_i}{p} + \left(\frac{n^2 - \sum x_i}{1-p}\right)(-1) = 0$$

$$\frac{\sum x_i}{p} = \left(\frac{n^2 - \sum x_i}{1-p}\right)$$

$$\sum x_i - \cancel{p\sum x_i} = n^2 p - \cancel{p\sum x_i}$$

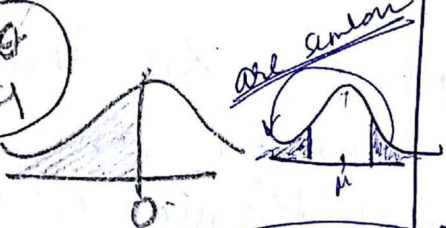
$$\boxed{\hat{p} = \frac{\sum x_i}{n^2}}$$

(c) $X_1 X_2 \ldots X_n \sim N(\theta, 1) \longrightarrow$ MLE of $\hat{\theta} = \bar{X}$

let $\delta = E\left[I_{X_1 > 0}\right]$ $\quad g(\theta) \to \text{MLE}[E[g(\theta)]]$

MLE of $\delta = ?$

$$\delta = E\left[I_{X_1 > 0}\right]$$
$$= P(X_1 > 0)$$
$$= 1 - P(X_1 \leq 0)$$
$$= 1 - F_{X_1}(0)$$
$$= 1 - \phi\left(\frac{0 - \mu}{\sigma}\right) \quad \mu = \theta,\ \sigma = 1$$



$$= \phi\left(\frac{\mu}{\sigma}\right)$$
$$= \phi\left(\frac{\theta}{1}\right) \quad \Rightarrow \text{MLE of } \delta = \phi(\bar{X})$$
$$= \phi\left(\frac{\sum X_i}{n}\right)$$
$$= \phi(\theta)$$

Scanned by CamScanner

6. Hypothesis Testing for a single population **(Total 13 points)**

(a) Consider the following 10 samples: {2.87, 1.09, 2.01, 0.93, 0.02, 1.78, 2.33, 0.65, 1.50, 0.99}. Use the K-S test to check whether these samples are from the Uniform(0, 3) distribution or not. First, set up the hypotheses. Then, create a 10 X 6 table with entries: $[x, \hat{F}_{-}(x), \hat{F}_{+}(x), F_0(x), |\hat{F}_{-}(x) - F_0(x)|, |\hat{F}_{+}(x) - F_0(x)|]$, where $\hat{F}_{-}(x)$ and $\hat{F}_{+}(x)$ are the values of the empirical distribution function to the left and right of x, and $F_0(x)$ is the CDF of Uniform(0, 3) at x. Finally, compare the max of the last two columns with the α = 0.05 threshold of 0.41 to Reject/Accept. (5 points)

(b) Assuming that the 10 samples are normally distributed, use the t-test to decide the null hypothesis that the population mean is 1.5. Use the α = 0.05 threshold of 2.228 to Reject/Accept. (3 points)

(c) You observe 46 successes in 100 trials of a coin. If the null hypothesis is that the coin is unbiased, use the Wald's test with the MLE/MME with α = 0.05 to Reject/Accept the null. What is the p-value? What if the null hypothesis is that the coin has p=0.7? What is the p-value? (5 points)

(a) Uniform (0,3) → pdf = $\frac{1}{3}$

→ $P(X \le x) = \frac{x}{3}$ for $0 < x < 3$

$H_0 : F(x) = F_0(x)$

$H_1 : F(x) \neq F_0(x)$ → CDF of Uniform (0,3)

↓ unknown CDF from which our data is sampled

→ critical value = 0.41 for α = 0.05,

⇒ cannot reject null Hypothesis

∴ 0.41 > 0.137 ~~crit~~ max value

(b) 10 samples → Normally dist

$H_0 : F(x) = F_0(x)$

$H_1 : F(x) \neq F_0(x)$ → CDF of Normal with $\mu = 1.5$

Sample Mean $\bar{X} = \frac{\sum x_i}{n} = \frac{14.17}{10} = 1.417$

Sample Var $\sigma^2 = \frac{\sum (x_i - \mu)^2}{n} = 0.680141$

t-test is

$$T = \frac{(\bar{X} - \mu)\sqrt{n}}{\hat{\sigma}_n} = \frac{(1.417 - 1.5)\sqrt{10}}{\sqrt{0.650141}} = -0.3$$

for $\alpha = 0.05$

critical value $= 2.228$

$$|T| = 0.326 < t_{n-1,\alpha} = 2.228$$

$\Rightarrow$ cannot reject claim that data was sampled from Normal dist

(c) $H_0: p = \frac{1}{2}$ $\quad$ size α Wald test is

$H_1: p \neq \frac{1}{2}$ $\quad$ reject H_0 when $|W| > z_{\alpha/2}$

where $\boxed{W = \frac{\hat{\theta} - \theta_0}{\hat{se}}}$

$X_1 X_2 \ldots X_n \sim \text{Bernoulli}(p)$

$$\hat{p} = \frac{\sum x_i}{n} = \frac{46}{100} = 0.46 \; ; \; \hat{\sigma}^2 = P(1-p)$$

$$= 0.46(1-0.46)$$

$$\Rightarrow W = \frac{\hat{\theta} - \theta_0}{(\hat{\sigma}/\sqrt{n})} = \frac{0.46 - 0.5}{\sqrt{\frac{0.46(1-0.46)}{100}}} = -0.8028$$

W

$$\text{p-value} = P(|W| > 0.8025)$$

$$\doteq 2P(W < -0.8026)$$

$$= 2 \times 0.2033 = 0.4066 > 0.05$$

$\Rightarrow$ cannot reject null Hyp

Now, consider,

$H_0 : p = 0.7$

$H_1 : p \neq 0.7$

then, $W = \frac{\hat{\theta} - \theta_0}{\hat{se}} = \frac{0.46 - 0.7}{\sqrt{\frac{0.46(1-0.46)}{100}}} = -4.818$

p-value $= P(|W| > 4.82)$

$= 2P(W < -4.82)$

$\approx 0 < 0.05$

$\Rightarrow$ We should reject null Hypothesis