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)

Amos Paul, Kanishta Agarwal, Prasad Marla, Sudeshna Pal

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

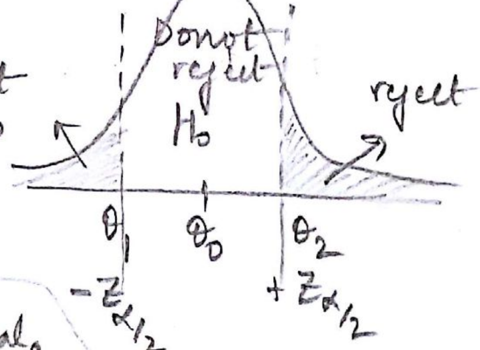
Suppose the null hypothesis is $H_0: \theta = \theta_0$, but the true value of θ is θ_*. Show that, under Wald's test, the probability of a Type II error is $\varphi(\frac{\theta_0-\theta_*}{\widehat{se}} + z_{\alpha/2}) - \varphi(\frac{\theta_0-\theta_*}{\widehat{se}} - z_{\alpha/2})$.

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

P(Type 2 error)

$= P(\text{Donot reject } H_0 \mid \theta = \theta_*)$

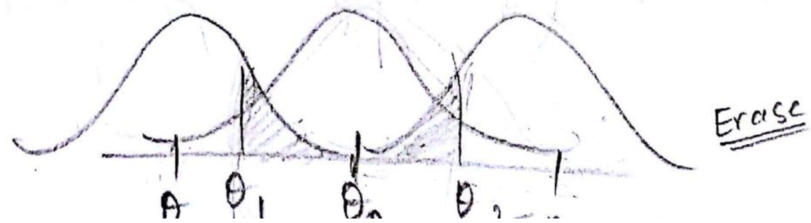
$= P(\theta < \theta_2 \mid \theta = \theta_*) - P(\theta < \theta_1 \mid \theta = \theta_*) \longrightarrow ①$



But, from definition of standard normal,

$$-z_{\alpha/2} = \frac{\theta_1 - \theta_0}{\widehat{se}} \Rightarrow \theta_1 = \theta_0 - \widehat{se}\, z_{\alpha/2}$$

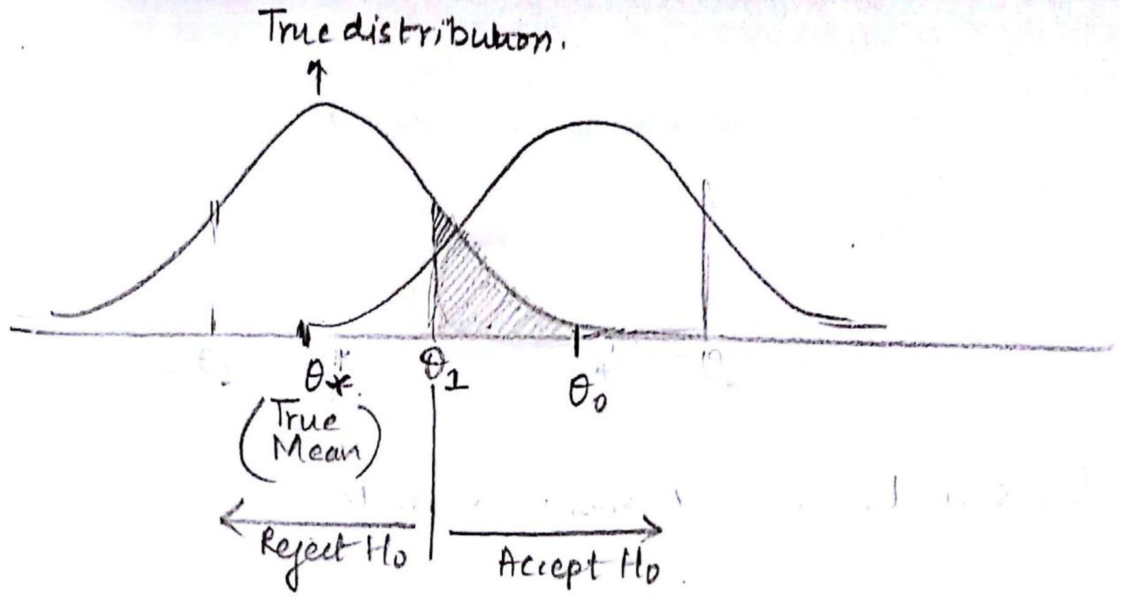
$$z_{\alpha/2} = \frac{\theta_2 - \theta_0}{\widehat{se}} \Rightarrow \theta_2 = \theta_0 + \widehat{se}\, z_{\alpha/2}$$



Erase

P.T.O

Scanned by CamScanner

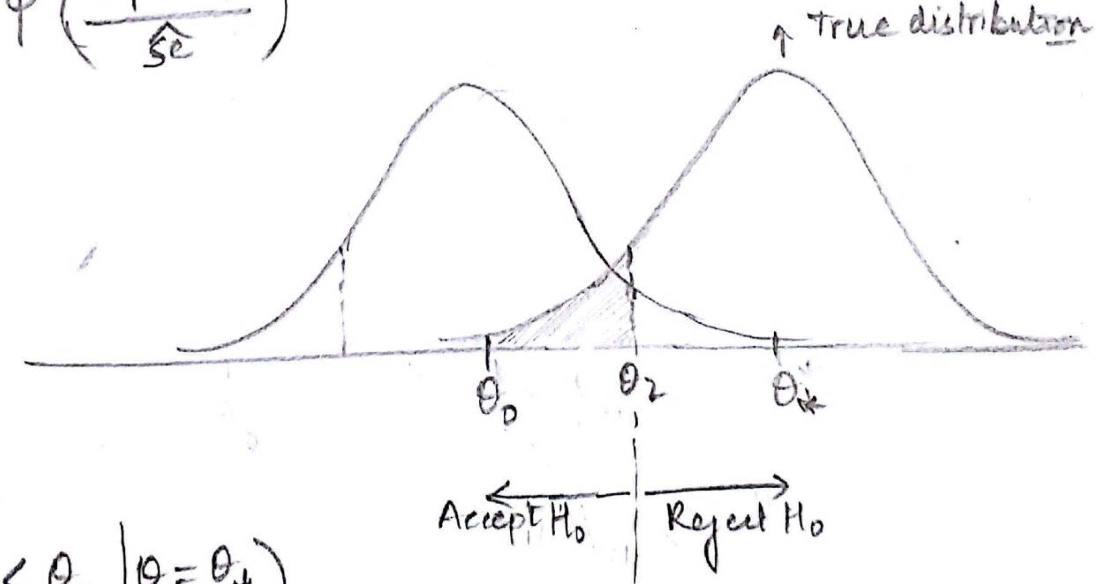


So, $P(\theta < \theta_1 \mid \theta = \theta_*)$

$$= P\left(Z_{\beta/2} < \frac{\theta_1 - \theta_*}{\hat{se}} \right) \quad \left(\text{from std normal distribution with true mean } \theta_* \right)$$

$$= \phi\left(\frac{\theta_1 - \theta_*}{\hat{se}} \right)$$

Illarly,



So, $P(\theta < \theta_2 \mid \theta = \theta_*)$

$$= P\left(Z_{\beta/2} < \frac{\theta_2 - \theta_*}{\hat{se}} \right) = \phi\left(\frac{\theta_2 - \theta_*}{\hat{se}} \right)$$

So, from eq ① & substituting values of θ_1 & θ_2

$$P(\text{type 2 error}) = \phi\left(\frac{\theta_0 + \hat{se}\, Z_{\alpha/2} - \theta_*}{\hat{se}} \right) - \phi\left(\frac{\theta_0 - \hat{se}\, Z_{\alpha/2} - \theta_*}{\hat{se}} \right)$$

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

Hence proved

Scanned by CamScanner

2. Posterior for Normal **(Total 10 points)**

Let $X_1, X_2, \ldots, X_n$ be distributed as Normal(θ, σ^2), where σ is assumed to be known. You are also given that the prior for θ is Normal(a, b^2).

(a) Show that the posterior of θ is Normal(x, y^2), such that: (7 points)

$x = \frac{b^2\bar{X}+se^2 a}{b^2+se^2}$ *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 $se^2 = \sigma^2/n$.

(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) Prior of θ = Normal(a, b^2)

from the definition of posterior of θ

$$f(\theta|x) \propto f(x|\theta) f(\theta)$$

Posterior $\propto$ Likelihood $\times$ prior

$$\propto L_n(\theta) f(\theta)$$

$$\Rightarrow N(x, y^2) \propto L_n(\theta) N(a, b^2)$$

$$f(x|\theta) = L_n(\theta) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_i-\theta)^2}{2\sigma^2}}$$

$$\Rightarrow N(x, y^2) \propto \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \exp\left\{-\sum_{i=1}^{n}\left(\frac{x_i^2+\theta^2-2\theta x_i}{2\sigma^2}\right)\right\} \frac{1}{\sqrt{2\pi b^2}} \exp\left\{-\frac{(\theta-a)^2}{2b^2}\right\}$$

$$\propto \frac{1}{(2\pi)^{\frac{n+1}{2}}(\sigma^n b)} \exp\left\{\frac{-\theta^2+2\theta a-a^2}{2b^2} - \sum_{i=1}^{n}\frac{x_i^2+\theta^2-2\theta x_i}{2\sigma^2}\right\}$$

As, we know σ and b ; we can ignore the constant term as it won't affect the proportionality constant.

$$\Rightarrow N(x, y^2) \propto \exp\left\{\frac{-\theta^2(\sigma^2+nb^2)+2\theta\left(a\sigma^2+b^2\sum_{i=1}^{n} x_i\right)-\left(a^2\sigma^2+b^2\sum_{i=1}^{n} x_i^2\right)}{2b^2\sigma^2}\right\}$$

Using $\frac{\sigma^2}{n} = se^2$ and $\bar{X} = \sum_{i=1}^{n} x_i$

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

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

$$\exp\left\{ -\frac{\left(a\,se^2 + \frac{b^2}{n}\Sigma x_i^2\right)}{b^2 se^2} \right\}$$

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

$$\Rightarrow x = \left(\frac{b^2\bar{x} + se^2 a}{b^2+se^2}\right) \; ; \; y = \left(\frac{b^2 se^2}{b^2+se^2}\right)$$

(b) Since, to find $(1-\alpha)$ posterior interval for θ

So, $N(x, y^2) \Rightarrow N(0,1)$

$$\Rightarrow Z = \left(\frac{\theta - \frac{b^2\bar{x} + se^2 a}{b^2+se^2}}{\frac{b^2 se^2}{b^2+se^2}}\right)$$

$$\Rightarrow P\left(-z_{\alpha/2} \leq \left(\frac{\theta - \frac{b^2\bar{x} + se^2 a}{b^2+se^2}}{\sqrt{\frac{b^2 se^2}{b^2+se^2}}}\right) \leq z_{\alpha/2}\right) = \alpha$$

$$P\left(-z_{\alpha/2}\sqrt{\left(\frac{b^2 se^2}{b^2+se^2}\right)} \leq \left(\theta - \frac{b^2\bar{x} + se^2 a}{b^2+se^2}\right) \leq z_{\alpha/2}\sqrt{\left(\frac{b^2 se^2}{b^2+se^2}\right)}\right) = \alpha$$

as $\sqrt{\left(\frac{b^2 se^2}{b^2+se^2}\right)}$ is positive,

$$P\left(-Z_{\alpha/2}\sqrt{\frac{b^2 se^2}{b^2+se^2}}\right)+\frac{b^2\bar{X}+se^2 a}{b^2+se^2} \le \theta \le \frac{b^2\bar{x}+se^2 a}{b^2+se^2}+Z_{\alpha/2}\sqrt{\frac{b^2 se^2}{b^2+se^2}}\Bigg)$$

$$= \alpha$$

$$\Rightarrow P\left(\frac{-Z_{\alpha/2}\, bse}{\sqrt{b^2+se^2}}+\frac{b^2\bar{x}+se^2 a}{b^2+se^2} \le \theta \le \frac{Z_{\alpha/2}\, bse}{\sqrt{b^2+se^2}}+\frac{b^2\bar{x}+se^2 a}{b^2+se^2}\right)=\alpha$$

$\Rightarrow (1-\alpha)$ Posterior interval of θ is.

$$\left(-\infty,\ \frac{-Z_{\alpha/2}\, b\cdot se}{\sqrt{b^2+se^2}}+\frac{b^2\bar{x}+se^2 a}{b^2+se^2}\right) \cup \left(\frac{Z_{\alpha/2}\, b\cdot se}{\sqrt{b^2+se^2}}+\frac{b^2\bar{x}+se^2 a}{b^2+se^2},\ \infty\right)$$

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.

$X_1 X_2 X_3 \dots X_n$ in Poisson(λ)

$$f(X=x) = \text{Poisson}(x;\lambda) = \frac{e^{-\lambda}\lambda^{x}}{\lfloor x} \Rightarrow L(\lambda;x) = \prod_{i=1}^{n} \frac{e^{-\lambda}}{\lfloor x_i}$$

Gamma(α,β) = prior of λ with
PDF of gamma$(\alpha,\beta) \propto x^{\alpha-1} e^{-x\beta}$

We know by definitions

$$f(\theta|X_1 X_2 \dots X_n) \propto f(X_1 X_2 \dots X_n|\theta) f(\theta)$$

Posterior $\propto$ liklihood $\times$ prior.

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

Here θ is parameter λ

$$f(\lambda|X_1 \dots X_n) \propto L_n(\lambda) f(\lambda)$$

$$\propto \prod_{i=1}^{n}\left(\frac{e^{-\lambda}\lambda^{x_i}}{\lfloor x_i}\right)\left(\lambda^{\alpha-1}e^{-\lambda\beta}\right)$$

$$\propto \frac{e^{-n\lambda}\lambda^{\sum_{i=1}^{n}x_i}}{\prod_{i=1}^{n}\lfloor x_i} \times \lambda^{\alpha-1} e^{-\lambda\beta}$$

$$\propto \lambda^{\sum x_i + \alpha - 1} \times e^{-(\beta+n)\lambda}$$

(You are calculating wrt λ
So, ignoring constant term $\frac{1}{\prod_{i=1}^{n}\lfloor x_i}$)

$$= \text{Gamma}\left(\sum X_i + \alpha, \beta + n\right) = \text{Gamma}(n\bar{X} + \alpha, \beta + n)$$

Hence posterior is also Gamma

So, if with prior on λ with $\lambda \sim$ Gamma(α, β)
and $X_i \sim$ Poisson(λ)

then, posterior on λ is $\sim \boxed{\text{Gamma}(n\bar{x}+\alpha, \beta+n)}$

So, posterior of Gamma(α, β) is a Gamma dist
with parameters
$n\bar{x}+\alpha$,
and $\beta+n$

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

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

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

(c) Let $X_1, X_2, \dots, X_n \sim$ 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) $X_1 X_2 X_3 \dots X_n$ are distributed as Poisson(λ)

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

Liklihood estimate } $L_n(\theta) = \prod_{i=1}^{n} \frac{\lambda^{x_i} e^{-\lambda}}{\lfloor x_i} = \frac{\lambda^{\sum_{i=1}^{n} x_i} e^{-n\lambda}}{\prod_{i=1}^{n} \lfloor x_i}$

log liklihood $\log(L_n(\theta)) = -n\lambda + \sum_{i=1}^{n} x_i \log\lambda - \log\left(\prod_{i=1}^{n} \lfloor x_i\right)$

find max; $\frac{d}{d\lambda} \log L_n(\theta) = -n + \sum_{i=1}^{n} x_i \frac{1}{\lambda} = 0$

$$\boxed{\hat{\lambda} = \frac{n}{\sum_{i=1}^{n} x_i}}$$

(b) $X_1 X_2 \dots X_m$ are distibuted as Binomial (n, p).

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

liklihood estimate $L_m(\theta) = \prod_{i=1}^{m} {}^{n}C_{x_i}\, p^{x_i} (1-p)^{n-x_i} = \left(\prod_{i=1}^{m} {}^{n}C_{x_i}\right) \cdot p^{\sum_{i=1}^{m} x_i} (1-p)^{mn-\sum_{i=1}^{m} x_i}$

log liklihood → $\log(L_m(\theta)) = \log\left(\prod_{i=1}^{m} {}^{n}C_{x_i}\right) + \sum_{i=1}^{m} x_i \log p + \left(mn - \sum_{i=1}^{m} x_i\right) \log(1-p)$

Scanned by CamScanner

Finding max;

So, $$\frac{d}{dp}\log\left(L_m(\theta)\right) = 0 + \frac{\sum_{i=1}^{m} x_i}{p} - \left(\frac{mn - \sum_{i=1}^{m} x_i}{1-p}\right) = 0$$

$$(1-p)\sum_{i=1}^{m} x_i - p\left(mn - \sum_{i=1}^{m} x_i\right) = 0$$

$$\sum_{i=1}^{m} x_i = p \cdot mn$$

$$\Rightarrow \boxed{\hat{p} = \frac{\sum_{i=1}^{m} x_i}{mn}}$$

(c) $$\delta = f(X) = E[I_{X>0}]$$

$$= P(X>0)$$

$$= \phi(X)$$

Now according to the invariance property

$$MLE(f(X)) = f(MLE(X)) \quad — (1)$$

Now $MLE(X) = \frac{1}{n}\sum_{i=1}^{n} X_i$

Substituting in (1)

$$MLE(\delta) = MLE(f(X)) = f(MLE(X))$$

$$= \phi\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) \;//$$