STAT245 Final Exercise Solution

1 Q1

(a) χ^2_n

(b) χ^2_{n-1}

(c) χ^2_{n-1}, a special case of (b) with $\mu = 0$

(d) χ^2_n since $(X_i - Y_i)/\sqrt{2} \sim N(0, \sigma^2)$ independently for $i = 1, ..., n$

(e) χ^2_{n+m-2}, see midterm question 1

(f) χ^2_2 since it's just $X^2 + Y^2$. Another way to derive χ^2_2 is to observe that $\frac{X-Y}{\sqrt{2}}$ and $\frac{X+Y}{\sqrt{2}}$ are independent $N(0, 1)$.

2 Q2

(a) CI for p with independent observations $X_1, ..., X_n \sim \text{Bernoulli}(p)$.

Proof.

$$\hat{p} = \frac{\sum_{i=1}^n X_i}{n} \sim \frac{1}{n} \sum_{i=1}^n \text{Bernoulli}(p)$$

$$\sqrt{n}(\hat{p} - p) \xrightarrow{D} N(0, p(1-p))$$

$$\left| \sqrt{n} \frac{\hat{p} - p}{\sqrt{p(1-p)}} \right| \leq z_{1-\alpha/2}$$

$$\text{CI for } p\text{: } \frac{(\hat{p} + \frac{z^2_{\alpha/2}}{2n}) \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z^2_{\alpha/2}}{4n^2}}}{1 + \frac{z^2_{\alpha/2}}{n}}$$

$\square$

(b) CI for λ with independent observations $X_1, ..., X_n \sim \text{Poisson}(\lambda)$.

Proof.

$$\hat{\lambda} = \frac{\sum_{i=1}^{n} X_i}{n} \sim \frac{1}{n}\sum_{i=1}^{n} \text{Poisson}(\lambda)$$

$$\sqrt{n}(\hat{\lambda} - \lambda) \xrightarrow{D} N(0, \lambda)$$

$$\left|\sqrt{n}\frac{\hat{\lambda} - \lambda}{\sqrt{\lambda}}\right| \leq z_{1-\alpha/2}$$

$$\text{CI for } \lambda\text{: } (\hat{\lambda} + \frac{z_{\alpha}^2}{2n}) \pm z_{\alpha/2}\sqrt{\frac{\hat{\lambda}}{n} + \frac{z_{\alpha/2}^2}{4n^2}}$$

□

(c) CI for σ^2 with independent observations $X_1, ..., X_n \sim N(\mu, \sigma^2)$.

Proof. Since $\sum_{i=1}^{n}(X_i - \overline{X})^2/\sigma^2 \sim \chi^2_{n-1}$, we know that

$$\sum_{i=1}^{n}(X_i - \overline{X})^2/\sigma^2 \stackrel{D}{=} \sum_{i=1}^{n-1} Z_i^2,$$

with $Z_1, ..., Z_{n-1} \sim N(0, 1)$ independently. Thus,

$$\frac{\hat{\sigma}^2}{\sigma^2} = \frac{1}{\sigma^2(n-1)}\sum_{i=1}^{n}(X_i - \overline{X})^2 = \frac{1}{n-1}\sum_{i=1}^{n-1} Z_i^2$$

$$\sqrt{n-1}\left(\frac{\hat{\sigma}^2}{\sigma^2} - 1\right) \xrightarrow{D} N(0, 2)$$

$$\left|\sqrt{\frac{n-1}{2}}\left(\frac{\hat{\sigma}^2}{\sigma^2} - 1\right)\right| \leq z_{1-\alpha/2}$$

$$\text{CI for } \sigma^2\text{: } \frac{\hat{\sigma}^2}{1 \pm \sqrt{\frac{2}{n-1}} z_{\alpha/2}}$$

□

(d) CI for σ^2 with $y \sim N(X\beta, \sigma^2 I_n)$.

Proof. Since $\|y - X\hat{\beta}\|^2/\sigma^2 \sim \chi^2_{n-p}$, we know that

$$\|y - X\hat{\beta}\|^2/\sigma^2 \stackrel{D}{=} \sum_{i=1}^{n-p} Z_i^2,$$

with $Z_1, ..., Z_{n-p} \sim N(0,1)$ independently. Thus,

$$\frac{\hat{\sigma}^2}{\sigma^2} = \frac{1}{\sigma^2(n-p)}\|y - X\hat{\beta}\|^2 = \frac{1}{n-p}\sum_{i=1}^{n-p} Z_i^2$$

$$\sqrt{n-p}\left(\frac{\hat{\sigma}^2}{\sigma^2} - 1\right) \xrightarrow{D} N(0,2)$$

$$\left|\sqrt{\frac{n-p}{2}}\left(\frac{\hat{\sigma}^2}{\sigma^2} - 1\right)\right| \leq z_{1-\alpha/2}$$

$$\text{CI for } \sigma^2 : \frac{\hat{\sigma}^2}{1 \pm \sqrt{\frac{2}{n-p}} z_{\alpha/2}}$$

$\square$

3 Q3

(a) see homework 2 question 2.

$$[\sin^2(\frac{2arcsin(\sqrt{\bar{x}}) + z_{\alpha/2}/\sqrt{n}}{2}), \sin^2(\frac{2arcsin(\sqrt{\bar{x}}) + z_{1-\alpha/2}/\sqrt{n}}{2})]$$

(b) see first week lecture notes

$$[(\sqrt{\bar{x}} - \frac{z_{1-\alpha/2}}{2\sqrt{n}})^2, (\sqrt{\bar{x}} + \frac{z_{1-\alpha/2}}{2\sqrt{n}})^2]$$

(c) See homework 6 question 1, just change $n - p$ to $n - 1$

$$[\hat{\sigma}^2 \exp(\sqrt{\frac{2}{n-1}} z_{\alpha/2}), \hat{\sigma}^2 \exp(\sqrt{\frac{2}{n-1}} z_{1-\alpha/2}])$$

(d) See homework 6 question 1.

$$[\hat{\sigma}^2 \exp(\sqrt{\frac{2}{n-p}} z_{\alpha/2}), \hat{\sigma}^2 \exp(\sqrt{\frac{2}{n-p}} z_{1-\alpha/2}])$$

4 Q4

(a) Find the MLE of β.

Proof.

$$p(Y) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{1}{2\sigma^2}(y_i - \beta x_i)^2\}$$
$$\frac{\partial \log p(Y)}{\partial \beta} = -\frac{1}{2\sigma^2} \sum_{i=1}^{n} x_i(\beta x_i - y_i)$$
$$\hat{\beta} = \frac{\sum_{i=1}^n x_i y_i}{\sum_{i=1}^n x_i^2}$$

□

(b) Find the distribution of the MLE $\hat{\beta}$.

Proof.

$$\hat{\beta} \sim \frac{1}{\sum_{i=1}^n x_i^2} \sum_{i=1}^{n} N(\beta x_i^2, x_i^2\sigma^2) \sim N\left(\beta, \frac{\sigma^2}{\sum_{i=1}^n x_i^2}\right)$$

□

(c) Show $\hat{\beta}$ and $\sum_{i=1}^n (y_i - \hat{\beta} x_i)^2$ are independent.

Proof.

$$\mathrm{Cov}(\hat{\beta}, y_1 - \hat{\beta} x_1) = \frac{\sum_{i=2}^n x_i^2}{\sum_{i=1}^n x_i^2} \cdot \frac{x_1}{\sum_{i=1}^n x_i^2} \cdot \mathrm{Var}(y_1) - \sum_{i=2}^{n} \frac{x_1 x_i}{\sum_{i=1}^n x_i^2} \cdot \frac{x_i}{\sum_{i=1}^n x_i^2} \cdot \mathrm{Var}(y_i) = 0.$$

By symmetry,

$$\mathrm{Cov}(\hat{\beta}, y_i - \hat{\beta} x_i) = 0$$

for all $i = 1, ..., n$. Then, $\hat{\beta}$ is independent of $y_i - \hat{\beta} x_i$ for all i, and thus the conclusion follows. □

(d) Find the distribution of $\sum_{i=1}^n (y_i - \hat{\beta} x_i)^2/\sigma^2$.

Proof. χ^2_{n-1}. □

(e) Construct a confidence interval for $\hat{\beta}$ using t-distribution.

Proof.

$$\hat{\sigma}^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \hat{\beta}x_i)^2 \sim \frac{\sigma^2}{n-1}\chi^2_{n-1}$$

$$\frac{(\sum_{i=1}^n x_i^2)^{1/2}(\hat{\beta}-\beta)/\sigma}{\hat{\sigma}/\sigma} \sim t_{n-1}$$

$$\text{CI for } \beta\colon \hat{\beta} \pm \frac{\hat{\sigma}}{(\sum_{i=1}^n x_i^2)^{1/2}} t_{n-1,\alpha/2}$$

□

(f) Consider prior distribution $\beta \sim N(0, \tau^2)$, find the posterior distribution of $\beta|y_1, ..., y_n$.

Proof.

$$\begin{aligned} p(\beta|y_1,\ldots,y_n) &\propto p(y_1,y_2,\ldots,y_n|\beta)p(\beta) \\ &\propto \prod_{i=1}^{n} \exp\left\{-\frac{(y_i-\beta x_i)^2}{2\sigma^2}\right\} \cdot \exp\left\{-\frac{\beta^2}{2\tau^2}\right\} \\ &\propto \exp\left\{-\frac{1}{2}\left(\frac{1}{\tau^2}+\frac{\sum_{i=1}^n x_i^2}{\sigma^2}\right)\left(\beta - \frac{\sum_{i=1}^n x_i y_i}{\frac{\sigma^2}{\tau^2}+\sum_{i=1}^n x_i^2}\right)^2\right\} \end{aligned}$$

Therefore $\beta|y_1, y_2, \ldots, y_n$ follows the normal distribution

$$N\left(\frac{\sum_{i=1}^n x_i y_i}{\frac{\sigma^2}{\tau^2}+\sum_{i=1}^n x_i^2}, \frac{1}{\frac{1}{\tau^2}+\frac{\sum_{i=1}^n x_i^2}{\sigma^2}}\right).$$

□

(g) Find the posterior mean.

Proof.

$$\mathbb{E}[\beta|y_1, y_2, \ldots, y_n] = \frac{\sum_{i=1}^n x_i y_i}{\frac{\sigma^2}{\tau^2}+\sum_{i=1}^n x_i^2}$$

□