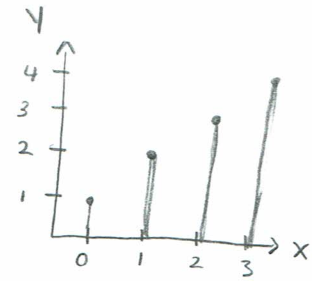
Ann Marie W.
Bio Theory Exam, 2016

1. Given $P(X=x) = (x+1)\,p^2 q^x$, $x = 0,1,2,\ldots$

$0<p<1$, prob of success

$q = 1-p$




a) Find joint density of X and Y.

$Y|X \sim Unif(0, x+1) \equiv \dfrac{1}{(x+1)} I(0<y<x+1)$

Then, $f_{X,Y}(x,y) = f_X(x)\cdot f_{Y|X}(y|x)$

$$= \cancel{(x+1)}\,p^2 q^x \frac{1}{\cancel{(x+1)}} I(0<y<x+1)$$

$$= \boxed{p^2 q^x I(0<y<x+1) \text{ for } \{(x,y) : x = \{0,1,2,\ldots\} \text{ and } 0<y<x+1\}}$$

b) Find $f_Y(y)$, the density of Y.

$$f_Y(y) = \sum_{x \in A(y)} f_{X,Y}(x,y) = \sum_{x\in A(y)} p^2 q^x I(0<y<x+1)$$

where $A(y) = \{\lfloor y \rfloor, \lfloor y \rfloor + 1, \ldots\}$

The $f_Y(y) = \displaystyle\sum_{x=\lfloor y\rfloor}^{\infty} p^2 q^x I(0<y<x+1)$

(Note: Remember formula for geometric series: $\displaystyle\sum_{n=M}^{\infty} cr^n = \frac{cr^M}{1-r}$)

$$= p^2 \sum_{x=\lfloor y \rfloor}^{\infty} q^x I(0<y<x+1) = \frac{p^2 q^{\lfloor y \rfloor}}{\underbrace{1-q}_{p}} = \boxed{p\, q^{\lfloor y \rfloor}, \quad 0<y<\infty}$$

Ann Marie W
Bios Theory Exam, 2016

1 c) Find $E[Y]$

Take $E[\underbrace{E[Y|X]}_{\sim \text{Unif}(0,\,x+1)}] = E\left[\frac{X+1}{2}\right] = \frac{1}{2}\underbrace{E(X)}_{\sim \text{Nbinom}(2,\,p)} + \frac{1}{2} = \frac{1}{\cancel{2}}\left[\frac{\cancel{2}(1-p)}{p}\right] + \frac{1}{2}$

$= \boxed{q/p + 1/2}$

d) Find $Cov(X, Y)$

Use double expectation formula for covariance

$$Cov(X,Y) = E[\underbrace{Cov(X, Y|X)}_{=0}] + Cov[E(Y|X), E(X|X)]$$

$$= \cancel{E[0]}^{\,0} + Cov\left[\frac{X+1}{2}, X\right] = \frac{1}{2}\underbrace{Var(X)}_{\sim \text{Nbinom}(2,p)} = \frac{1}{2}\left[\frac{2q}{p^2}\right] = \boxed{\frac{q}{p^2}}$$

e) Define $T = 2Y - X$. Find $Cov(T, X)$

$$Cov(T,X) = Cov(2Y-X, X) = 2\,Cov(Y,X) - Cov(X,X)$$

$$= 2\left(\frac{q}{p^2}\right) - Var(X) = \frac{2q}{p^2} - \frac{2q}{p^2} = 0.$$

f) Are T and X independent?

Hint: On these types of problems, they will try to trick you into saying that $Cov(T,X) = 0 \Rightarrow T \perp\!\!\!\perp X$. However, this is NOT true unless T and X are jointly normal.

Take $P(2Y - X = 2Y \,|\, X = 0) = 1 \neq P(2Y - X = 2Y)$.

Thus, $T = 2Y - X$ is not independent from X.

Ann Marie W
Bios Theory Exam, 2016

2. $X_1, \ldots, X_n \sim N(\mu, 1)$

Define $\theta = P(X>0)$. Use $\Phi(t)$ to denote CDF of $N(0,1)$ evaluated @ t.

a) Express $P(X>0)$ as a function of $\mu$.

$$P(X>0) = 1 - P(X \le 0) = 1 - P\Big(\underbrace{\tfrac{X-\mu}{1}}_{\sim N(0,1)} \le -\mu\Big) = 1 - \Phi(-\mu) = \boxed{\Phi(\mu)}$$

b) Find an unbiased estimator of $P(X>0)$

Want an unbiased estimator of $\theta = P(X>0)$ (a.k.a. $E[\overset{\text{unbiased}}{\text{estimator}}] = \theta$).

Since the thing we want to estimate is a probability, this is a hint that the estimator is likely to be an indicator function.

Take $E[I(X>0)] = P(X>0)$. Thus, $I(X>0)$ is an unbiased estimator of $\theta = P(X>0)$.

c) Find the MLE of $P(X>0)$.

First, find MLE of $\mu$.

$$L(\mu|x) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum_{i=1}^n (x_i-\mu)^2/2} \Rightarrow \ell(\mu|x) = \log\left[\left(\frac{1}{\sqrt{2\pi}}\right)\right] - \sum_{i=1}^n (x_i-\mu)^2/2$$

$$\Rightarrow \frac{\partial \ell}{\partial \mu} = -\cancel{2}\sum_{i=1}^n (x_i-\mu)/\cancel{2} \overset{\text{set}}{=} 0 \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^n x_i = \bar{x}$$

Since $\frac{\partial^2 \ell}{\partial \mu^2} = -n < 0 \Rightarrow \hat{\mu}$ occurs @ a global max.

Then, since $P(X>0) = 1 - \Phi(-\mu)$, as in a), have $\hat{\theta} = \widehat{P(X>0)} = 1 - \Phi(-\hat{\mu})$

$= 1 - \Phi(-\bar{x}) = \Phi(\bar{x})$ by the invariance property of MLE.

Ann Marie W.
Bios Theory Exam, 2016

2 d) Find the Cramer-Rao Lower bound on the variance of unbiased estimators of $P(X>0)$.

To calculate CRLB, can use formula:

$$CRLB = \frac{\left\{\frac{d\tau(\theta)}{d\theta}\right\}^2}{-E\left\{\frac{\partial^2}{\partial\theta^2}\log f(\underset{\sim}{x}|\theta)\right\}}$$

where $\tau(\theta) = E(W)$ where $W$ is an unbiased estimator of $\theta$.

Here $\tau(\theta) = \theta \Rightarrow \left\{\frac{d\tau(\theta)}{d\theta}\right\}^2 = (1)^2 = 1$

$$-E\left\{\frac{\partial^2}{\partial\theta^2}\left(\log\left[\left(\frac{1}{\sqrt{2\pi}}\right)^n\right] - \sum_{i=1}^n (x_i-\theta)^2\right)\right\} = -E(-n) = n$$

Thus, $\boxed{CRLB = 1/n}$

2e) Find the UMVUE of $P(X>0)$.

First, show that $\sum_{i=1}^n X_i$ is a CSS for $\mu$.

Have $f(\underset{\sim}{x}, \mu) = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum_{i=1}^n (x_i-\mu)^2/2} = \left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum_{i=1}^n (x_i^2 - 2\mu x_i + \mu^2)/2}$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\sum_{i=1}^n x_i^2/2}}_{h(x)} \cdot e^{\overbrace{\sum_{i=1}^n x_i}^{t(x)} \underbrace{\mu}_{w(\mu)}} \cdot \underbrace{e^{-\sum_{i=1}^n \mu^2/2}}_{\underbrace{e^{-n\mu^2/2}}_{c(\mu)}}$$

where $\mu \in (-\infty, \infty)$, an open set in $\mathbb{R}^1$.

$\Rightarrow \sum_{i=1}^n X_i$ is a CSS for $\mu \Rightarrow \bar{X}$ is also a CSS for $\mu$.

Since $\bar{X}$ is a CSS for $\mu$ and since $\Phi(\bar{x})$ is the MLE of $\Phi(\mu)$, then $\Phi(\bar{X})$ is the UMVUE of $E(\Phi(\bar{X})) = \Phi(\mu) = P(X>0)$.

Ann Marie W.
Bios Theory Exam, 2016

3. Given $X_1 \in \{a_1, a_2, a_3, a_4\}$

| | $a_1$ | $a_2$ | $a_3$ | $a_4$ |
|---|---|---|---|---|
| $\theta_1$ | 0.3 | 0.4 | 0.1 | 0.2 |
| $\theta_2$ | 0.4 | 0.1 | 0.2 | 0.3 |
| $\theta_3$ | 0.2 | 0.1 | 0.5 | 0.2 |

Note: No summation in likelihood since given single obs, $X$.

a) Find the MLE of $\theta$ under different values of $X$

Given one observation, $X$, need to find the value $\theta_i$ at which the global max occurs.

$L(\theta|x) = P(a_i|\theta)$ for $i \in \{1,2,3\}$

From table, can see that
$$\hat{\theta} = \begin{cases} \theta_2, & x = a_1 \\ \theta_1, & x = a_2 \\ \theta_3, & x = a_3 \\ \theta_2, & x = a_4 \end{cases} = \boxed{\begin{cases} \theta_2, & x = a_1, a_4 \\ \theta_1, & x = a_2 \\ \theta_3, & x = a_3 \end{cases}}$$

b) Derive the critical region of the LRT for $H_0: \theta = \theta_1$ vs. $H_1: \theta \neq \theta_1$ with type I error prob $\alpha = 0.1$ and $\Theta = \{\theta_1, \theta_2, \theta_3\}$

$$\lambda(x) = \frac{\sup_{\theta \in \Theta_0} L(\theta|x)}{\sup_{\theta \in \Theta} L(\theta|x)} = \frac{\sup_{\theta \in \Theta_0} P(a_i|\theta)}{\sup_{\theta \in \Theta} P(a_i|\theta)}$$

$$= \begin{cases} \frac{P(a_1|\theta_1)}{P(a_1|\theta_2)} = \frac{0.3}{0.4} = \frac{3}{4}, & x = a_1 \\ \frac{P(a_2|\theta_1)}{P(a_2|\theta_1)} = \frac{0.4}{0.4} = 1, & x = a_2 \\ \frac{P(a_3|\theta_1)}{P(a_3|\theta_3)} = \frac{0.1}{0.5} = \frac{1}{5}, & x = a_3 \\ \frac{P(a_4|\theta_1)}{P(a_4|\theta_2)} = \frac{0.2}{0.3} = \frac{2}{3}, & x = a_4 \end{cases}$$

Then, the critical region $R = \{x : \lambda(x) \leq c\}$ for $c \in [0,1]$.

$\Rightarrow \underset{0.1}{\alpha} = \sup_{\theta \in \Theta_0} P(\lambda(x) \leq c)$

$= P(\lambda(x) \leq c \mid \theta = \theta_1)$

Then, $P(a_3|\theta_1) = 0.1 = P(\lambda(x) \leq 1/5)$

$\Rightarrow \boxed{R = \{x : \lambda(x) \leq 1/5\}}$

3 c) Give the test function of the LRT in b) in explicit form. Explain explicitly how one would apply the testing procedure given a single obs. X.

Have $\lambda(x) = \dfrac{\sup_{\theta \in \Theta_0} L(\theta|x)}{\sup_{\theta \in \Theta} L(\theta|x)}$ is the likelihood ratio test for testing

$H_0 : \theta \in \Theta_0$ against $H_1 : \theta \in \Theta_0^c$. Let $\hat{\theta}_0$ denote the restricted MLE over $\Theta_0$ and let $\hat{\theta}$ denote the unrestricted MLE over $\Theta = \Theta_0 \cup \Theta_0^c$.

Then, the LRT statistic is:

rejection region

$\lambda(x) = \dfrac{L(\theta_0|x)}{L(\hat{\theta}_{MLE}|x)}$. Then, $R = \{x : \lambda(x) \leq c\}$ for $c \in [0,1]$

$\Leftrightarrow R^* = \{x : \hat{\theta}_{MLE} \geq c^* \text{ or } \hat{\theta}_{MLE} \leq c^*\}$ where $\hat{\theta} \geq c^*$ or $\hat{\theta} \leq c^*$ follows the direction of $H_1$.

No clue if this is what they want.
The question isn't very "explicit." -pun intended

AMW
Bios Theory Exam, 2016

3 d) Find the UMP test for testing $H_0: \theta = \theta_1$ vs. $H_1: \theta = \theta_2$ w/ $\alpha = 0.1$

$$R = \left\{ x \in \{a_1, a_2, a_3, a_4\} : \frac{f(x|\theta_2)}{f(x|\theta_1)} > k \right\}$$

The ratios of pmfs give:

$$\frac{f(a_1|\theta_2)}{f(a_1|\theta_1)} = \frac{0.4}{0.3} = \frac{4}{3}, \qquad \frac{f(a_2|\theta_2)}{f(a_2|\theta_1)} = \frac{0.1}{0.4} = \frac{1}{4}$$

$$\frac{f(a_3|\theta_2)}{f(a_3|\theta_1)} = \frac{0.2}{0.1} = 2 \qquad \frac{f(a_4|\theta_2)}{f(a_4|\theta_1)} = \frac{0.3}{0.2} = \frac{3}{2}$$

If we choose $\frac{1}{4} < k < 2$, the Neyman-Pearson Lemma says that the test rejects $H_0$ if $x = a_3$ is the UMP level $\alpha = P(x = a_3|\theta_1) = 0.1$ test.

3e) Comment on whether the UMP test for the hypothesis in d) is also the UMP test for the hypothesis in b). If you think it is, provide the rationale. If you think it is not, derive the UMP test for the hypothesis in b).

No, we cannot compare the UMP test in 3d) to the UMP test in 3b) because 3b) does not have a UMP test. For 3b), $H_0: \theta = \theta_1$ vs. $H_1: \theta \neq \theta_1$. Simply put, a UMP test does not exist for testing as given in 3b) because critical regions turn out to be different for $\theta > \theta_1$ and $\theta < \theta_1$. This means, there are only UMP tests for one-sided hypotheses in which we can use the N-P lemma.