Name: KEY

MATH 321: Review Part 2

1. Let X_1 and X_2 be independent random variables where $X_1 \sim \text{Normal}\,(\mu = 1, \sigma = 2)$, and $X_2 \sim \text{Normal}\,(\mu = 5, \sigma = 3)$.

   Find $P(3X_1 - X_2 < 2)$.

Let $Y = 3X_1 - X_2 \sim \text{Normal}\left(\mu = 3\mu_1 - \mu_2 = 3(1) - 5 = -2,\ \sigma^2 = 3^2\sigma_1^2 + \sigma_2^2 = 3^2 2^2 + 3^2 = 45\right)$

$\Rightarrow P(3X_1 - X_2 < 2) = P(Y < 2) = \text{normalcdf}\left(\text{lower} = -10000,\ \text{upper} = 2,\ \mu = -2,\ \sigma = \sqrt{45}\right) \approx \boxed{0.7245}$

2. Let $X_{(1)}, \ldots, X_{(8)}$ be the order statistics of a random sample of size $n = 8$ from a random variable X with pdf

$$f(x) = \frac{3}{4}(x^2 + 1), \quad 0 \le x \le 1$$

   (a) Find the cdf of X, $F_X(x)$, and $P(X \le 0.25)$.

$\rightarrow F_X(x) = P(X \le x) = \int_{-\infty}^{x} f(u)\,du = \int_0^x \frac{3}{4}(u^2+1)\,du = \frac{1}{4}u^3 + \frac{3}{4}u \Big|_0^x$

$\boxed{= \frac{1}{4}(x^3 + 3x), \ 0 \le x \le 1}$

$\rightarrow P(X \le 0.25) = F_X(0.25) = \frac{1}{4}(0.25^3 + 3(0.25))$

$\boxed{\approx 0.1914}$

   (b) Find the cdfs $F_{X_{(1)}}(x)$ and $F_{X_{(8)}}(x)$.

$\rightarrow F_{X_{(1)}}(x) = P(X_{(1)} \le x) = 1 - P(\text{ALL } X_i > x) = 1 - (1 - F_X(x))^8$ (complement)

$\boxed{= 1 - \left[1 - \frac{1}{4}(x^3 + 3x)\right]^8, \ 0 \le x \le 1}$

$\rightarrow F_{X_{(8)}}(x) = P(X_{(8)} \le x) = P(\text{ALL } X_i \le x) = (F_X(x))^8$

$\boxed{= \left[\frac{1}{4}(x^3 + 3x)\right]^8, \ 0 \le x \le 1}$

can skip to last steps & no NEED to simplify

   (c) Find the pdfs $f_{X_{(1)}}(x)$ and $f_{X_{(8)}}(x)$.

$f_{X_{(1)}}(x) = \frac{d}{dx} F_{X_{(1)}}(x) \rightarrow f_{X_{(1)}}(x) = \frac{d}{dx}\left[1 - \left[1 - \frac{1}{4}(x^3 + 3x)\right]^8\right]$

$\boxed{= +8\left[1 - \frac{1}{4}(x^3 + 3x)\right]^7 \left[\frac{3}{4}(x^2 + 1)\right], \ 0 \le x \le 1}$

OR

* shortcuts for extreme order stats $\rightarrow f_{X_{(1)}}(x) = n\left[1 - F_X(x)\right]^{n-1} f_X(x) =$

$\rightarrow f_{X_{(n)}}(x) = n\left[F_X(x)\right]^{n-1} f_X(x)$

$\boxed{= 8\left[\frac{1}{4}(x^3 + 3x)\right]^7 \left[\frac{3}{4}(x^2 + 1)\right], \ 0 \le x \le 1}$

(d) Find the cdf $F_{X_{(6)}}(x)$

$$\rightarrow F_{X_{(j)}}(x) = \sum_{k=j}^{n} \binom{n}{k} [F_X(x)]^k [1 - F_X(x)]^{n-k}$$

No NEED to simplify

$$\boxed{F_{X_{(6)}}(x) = \sum_{k=6}^{8} \binom{8}{k} \left[\frac{1}{4}(x^3 + 3x)\right]^k \left[1 - \frac{1}{4}(x^3 + 3x)\right]^{8-k}, \quad 0 \le x \le 1}$$

(e) Find $P(X_{(6)} \le 0.25)$.

$$P(X_{(6)} \le 0.25) = F_{X_{(6)}}(0.25) = \sum_{k=6}^{8} \binom{8}{k} [\underbrace{F_X(0.25)}_{=0.1914}]^k [1 - \underbrace{F_X(0.25)}_{=1-0.1914}]^{8-k} \approx 0.0009627$$

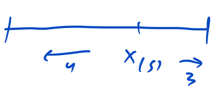
OR $\rightarrow$ If $Y = \#\ X_i \le 0.25$

$\downarrow \sim Bin(n=8, p=0.1914)$

$P(X_{(6)} \le 0.25) = P(\text{At least } 6\ X_i \le 0.25) = P(Y \ge 6) = 1 - \text{Binomcdf}(n=8, p=0.1914, x=5) = 0.0009627$

(f) Find the pdf $f_{X_{(5)}}(x)$.

$$\rightarrow f_{X_{(j)}}(x) = \frac{n!}{(j-1)!\,1!\,(n-j)!} [F_X(x)]^{j-1} f_X(x) [1 - F_X(x)]^{n-j}$$



$$\boxed{f_{X_{(5)}}(x) = \frac{8!}{4!\,1!\,3!} \left[\frac{1}{4}(x^3 + 3x)\right]^4 \frac{3}{4}(x^2 + 1) \left[1 - \frac{1}{4}(x^3 + 3x)\right]^3, \quad 0 \le x \le 1}$$

(g) Write the integrals to find $E(X_{(5)})$ and $P(X_{(5)} \ge 0.75)$

$$\boxed{E(X_{(5)}) = \int_{-\infty}^{\infty} x f_{X_{(5)}}(x)\,dx = \int_0^1 x \left[\frac{8!}{4!\,1!\,3!} \left[\frac{1}{4}(x^3 + 3x)\right]^4 \frac{3}{4}(x^2 + 1) \left[1 - \frac{1}{4}(x^3 + 3x)\right]^3\right] dx}$$

$$\boxed{P(X_{(5)} \ge 0.75) = \int_{0.75}^{\infty} f_{X_{(5)}}(x)\,dx = \int_{0.75}^{1} \frac{8!}{4!\,1!\,3!} \left[\frac{1}{4}(x^3 + 3x)\right]^4 \frac{3}{4}(x^2 + 1) \left[1 - \frac{1}{4}(x^3 + 3x)\right]^3 dx}$$

3. Let $X_1, \ldots, X_n$ be a random sample from $f(x \mid \theta) = \theta x, \quad 0 < x < 1,\ \theta > 0$.

(a) Find the MME of θ.

$\rightarrow E(X) = \int_{-\infty}^{\infty} x f(x)\,dx = \int_0^1 x(\theta x)\,dx = \frac{1}{3}\theta x^3 \Big|_0^1 = \frac{1}{3}\theta$

$\rightarrow \mu_1' = \frac{1}{3}\theta$

$\rightarrow m_1' = \bar{x}$

$\rightarrow \frac{1}{3}\theta = \bar{x} \Rightarrow \boxed{\hat{\theta}_{MME} = 3\bar{x}}$

(b) Show if $\hat{\theta}_{MME}$ is unbiased.

$E(\hat{\theta}_{MME}) = E(3\bar{x}) = 3\underbrace{E(\bar{x})}_{=\mu} = 3\left(\frac{1}{3}\theta\right) = \theta \Rightarrow$ unbiased ✓

4. Let $X_1, \ldots, X_n$ be a random sample from $f(x \mid \theta) = (\theta+1)x^\theta, \quad 0 < x < 1,\ \theta > -1$.

(a) Find the MLE of θ.

① $L(\theta|x) = \prod_{i=1}^{n} f(x_i|\theta) = \prod_{i=1}^{n} (\theta+1) x_i^\theta = (\theta+1)^n \prod_{i=1}^{n} x_i^\theta$

$\ell(\theta) = \ln[L(\theta|x)] = \ln[(\theta+1)^n \prod x_i^\theta] = \ln[(\theta+1)^n] + \ln[\prod x_i^\theta]$

$= n\ln(\theta+1) + \sum \ln(x_i^\theta)$

$= \quad + \sum \theta \ln(x_i)$

$= \quad + \theta \sum \ln(x_i)$

② $\rightarrow \ell'(\theta) = \frac{d}{d\theta}[n\ln(\theta+1) + \theta \sum \ln(x_i)]$

$= \frac{n}{\theta+1} + \sum \ln(x_i)$

$\rightarrow 0 = \frac{n}{\theta+1} + \sum \ln(x_i)$

$-\sum \ln(x_i) = \frac{n}{\theta+1}$

$-(\theta+1)\sum \ln(x_i) = n$

$-\theta \sum \ln(x_i) - \sum \ln(x_i) = n$

$\Rightarrow \hat{\theta} = \frac{n + \sum \ln(x_i)}{-\sum \ln(x_i)} = \frac{-n}{\sum \ln(x_i)} - 1$

③ $\ell''(\theta) = \frac{d}{d\theta}\left[\frac{n}{\theta+1} + \sum \ln(x_i)\right]$

$= \frac{-n}{(\theta+1)^2}$

$\ell''(\hat{\theta}) = \frac{-n}{\left(\frac{-n}{\sum \ln(x_i)} - 1 + 1\right)^2} < 0$ ✓

$\Rightarrow \boxed{\hat{\theta}_{MLE} = \frac{-n}{\sum \ln(x_i)} - 1}$

(b) Find the MLE of $E(X) = \dfrac{\theta+1}{\theta+2}$.

$E(X) = \hat{\tau}(\theta) = \tau(\hat{\theta}) = \dfrac{\left(\frac{-n}{\sum \ln(x_i)} - 1\right) + 1}{\left(\frac{-n}{\sum \ln(x_i)} - 1\right) + 2} = \boxed{\dfrac{\frac{-n}{\sum \ln(x_i)}}{1 - \frac{n}{\sum \ln(x_i)}}}$ extra simplifying $= \dfrac{\frac{-n}{\sum \ln(x_i)}}{\frac{\sum \ln(x_i) - n}{\sum \ln(x_i)}} = \boxed{\frac{-n}{\sum \ln(x) - n}}$

$E(x) = 1/p = 1/0.45 = 2.22$

$V(x) = q/p^2 = \frac{0.55}{0.45^2} = 2.716$

5. Let $X_1, \ldots, X_{35}$ be a random sample from a Geometric ($p = 0.45$) experiment.

   (a) Let Y_1 be the average number of trials to get the first success in the $n = 35$ runs of this experiment, i.e. $Y_1 = \bar{X}$. Find the distribution of Y_1.

$n = 45 \geq 30 \Rightarrow$ By CLT

$$\boxed{Y_1 = \bar{X} \approx Normal\left(\mu = E(x) = 1/0.45,\ \sigma^2 = \frac{V(x)}{n} = \frac{0.55}{35(0.45^2)}\right)}$$

   (b) Let Y_2 be the total number of runs (across all trials) to get the first success in $n = 35$ trials of this experiment, i.e. $Y_2 = \sum_{i=1}^{35} X_i$. Find the distribution of Y_2.

By CLT

$$\boxed{Y_2 = \sum_{i=1}^{35} X_i \approx Normal\left(\mu = nE(x) = 35(1/0.45),\ \sigma^2 = nV(x) = \frac{35(0.55)}{0.45^2}\right)}$$

exact distribution

Σ Geometric = Negative Binomial

$\Rightarrow Y \sim NB(r = 35, p = 0.45)$

6. Let $X_1, \ldots, X_{40}$ be a random sample from a $f(x)$, where $\mu = 15$ and σ^2 is unknown. Let $\bar{X}$ be the sample mean and S be the sample standard deviation.

   (a) Find the distribution of $\bar{X}$.

$n = 40 \geq 30 \Rightarrow$ By CLT

$$\boxed{\bar{X} \overset{approx}{\sim} Normal\left(\underset{\mu = E(X)}{\mu = 15},\ \underset{\sigma^2 = \frac{V(X)}{n}}{\frac{\sigma^2}{40}} \approx \frac{S^2}{40}\right)}$$

Substitute S^2 for unknown σ^2

   (b) Find the distribution of $\dfrac{\bar{X} - 15}{S/\sqrt{40}}$.

n is large

$$\boxed{\approx Z \overset{approx}{\sim} Normal(0,1)}$$

   (c) Suppose we know the population variance $\sigma^2 = 4$. Find the distribution of $\dfrac{\bar{X} - 15}{2/\sqrt{40}}$.

$$\boxed{= \frac{\bar{X} - \mu}{\sigma/\sqrt{n}} \approx Z}$$

b/c X_i not Normal, but Known σ^2 $\Rightarrow$ good approximation