CSE 312 Practice Final Solutions

1. True/False, Short Answer. Provide a short justification for your answer.

a) True or False. For any μ and σ^2, if $X \sim N(\mu, \sigma^2)$, then $Z = \frac{X-\mu}{\sigma^2} \sim N(0,1)$.

False. $Z = \frac{X-\mu}{\sigma} \sim N(0,1)$.

b) True or False. For any μ and σ^2, if $X \sim N(\mu, \sigma^2)$ has pdf $f_X(x)$, then $\int_0^\infty f_X(x)dx = 1/2$.

False. The normal density is symmetric around μ, not 0, except for the single case when $\mu = 0$.

c) Let X have a continuous uniform distribution, $X \sim Unif(a,b)$. What is $E[X^3]$, the skewness of X? Do not simplify further than evaluating the integral.

$E[X^3] = \int_a^b \frac{x^3}{b-a} dx = \frac{1}{b-a}\left[\frac{1}{4}x^4\right]_a^b = \frac{1/4}{b-a}(b^4 - a^4)$.

d) True or False. Suppose X has a continuous uniform distribution, $Unif(0,1)$. If $Y = -\ln X$, then $Y \sim Exp(1)$.

True. $P(X \le x) = x$, for any $x \in [0,1]$. $F_Y(y) = P(Y \le y) = P(-\ln X \le y) = P(X \ge e^{-y}) = 1 - e^{-y}$, for any $y \ge 0$. This cdf uniquely defines the $Exp(1)$ distribution.

e) True or False. For any random variable X and any $\alpha > 0$, $P(X \ge \alpha) \le \frac{E[X]}{\alpha}$.

False. This is almost Markov's Inequality, but Markov's requires that X is non-negative.

f) True or False. If $Z \sim N(0,1)$, then $P(Z \ge -z) = \Phi(z)$.

True. $P(Z \ge -z) = 1 - P(Z \le -z) = 1 - \left(1 - \Phi(z)\right) = \Phi(z)$.

g) True or False. For any random variable X, $P(X \le x) = P(X < x)$.

False. This is only true for continuous random variables.

h) True or False. If $X_1, X_2, \ldots, X_n$ are iid, each with standard deviation σ, then $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ has standard deviation $\sigma_{\bar{X}} = \sigma/n$.

False. $Var(\bar{X}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} n\sigma^2 = \frac{\sigma^2}{n} \to \sigma_{\bar{X}} = \frac{\sigma}{\sqrt{n}}$.

i) Suppose $Z \sim N(0,1)$. Find $Var(Var(Z^3))$.

$Var(Z^3)$ is a constant. $Var(constant) = 0$.

j) Suppose $X \sim N(\mu_1, \sigma_1^2)$. If we want to apply a linear transformation f to X to get $X' = f(X) \sim N(\mu_2, \sigma_2^2)$, what should $f(X)$ be?

$\frac{X-\mu_1}{\sigma_1} \sim N(0,1)$. $\frac{X'-\mu_2}{\sigma_2} \sim N(0,1)$. So $\frac{X-\mu_1}{\sigma_1} = \frac{X'-\mu_2}{\sigma_2} \to f(X) = X' = \sigma_2\left(\frac{X-\mu_1}{\sigma_1}\right) + \mu_2$.

k) True or False. Suppose X is a continuous random variable with pdf $f_X(x)$. Then, $0 \leq f_X(x) \leq 1 \; \forall x \in \mathbb{R}$.

False. Consider $X \sim Unif\left(0, \frac{1}{2}\right)$. Densities are not probabilities.

l) True or False. Suppose X is a discrete random variable with pmf $p_X(x)$. Then, $0 \leq p_X(x) \leq 1 \; \forall x \in \mathbb{R}$.

True. Probability mass functions are probabilities.

m) True or False. Suppose X is a continuous random variable with cdf $F_X(x)$. Then, $0 \leq F_X(x) \leq 1 \; \forall x \in \mathbb{R}$.

True. Cumulative distribution functions represent cumulative probabilities, which integrate to 1.

n) True or False. Suppose X is a continuous random variable with cdf $F_X(x)$. Then, $F_X(x)$ is a strictly increasing function. (That is, $\forall a, b \in \mathbb{R}, a < b \to F_X(b) > F_X(a)$).

False. Where there's 0 density, the cdf may remain constant.

o) True or False. Let $A_1, A_2, \ldots, A_n$ be events such that $A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$ and $A_1 \cap A_2 \cap \ldots \cap A_n = \emptyset$, and let B be any event. Then,

$$P(B) = P(A_1 \cap B) + \cdots + P(A_n \cap B)$$

False. $A_1, \ldots, A_n$ have to be pairwise mutually exclusive.

2. Alex decided he wanted to create a "new" type of distribution that will be famous, but he needs some help. He knows he wants it to be continuous and have uniform density, but he needs help working out some of the details. We'll denote a random variable X having the "Uniform-2" distribution as $X \sim Unif2(a, b, c, d)$, where $a < b < c < d$. We want the density to be non-zero in $[a, b]$ and $[c, d]$, and zero everywhere else. Anywhere the density is non-zero, it must be equal to the same constant.

a) Find the probability density function, $f_X(x)$. Be sure to specify the values it takes on for every point in $(-\infty, \infty)$. (Hint: use a piece-wise definition).

$$f_X(x) = \begin{cases} \dfrac{1}{(b-a)+(d-c)}, & x \in [a,b] \cup [c,d] \\ 0, & otherwise \end{cases}$$

b) Find the cumulative distribution function, $F_X(x)$. Be sure to specify the values it takes on for every point in $(-\infty, \infty)$. (Hint: use a piece-wise definition).

$$F_X(x) = \begin{cases} 0, & x \in (-\infty, a) \\ \dfrac{(x-a)}{(b-a)+(d-c)}, & x \in [a,b) \\ \dfrac{(b-a)}{(b-a)+(d-c)}, & x \in [b,c) \\ \dfrac{(b-a)+(x-c)}{(b-a)+(d-c)}, & x \in [c,d) \\ 1, & x \in [d, \infty) \end{cases}$$

c) [Extra Credit] Suppose $x_1, x_2, \ldots, x_n$ are iid from $Unif2(a, b, c, d)$. For simplicity, you may assume $n \geq 4$. Find the maximum likelihood estimators for a, b, c, and d. (Hint: Do not take any derivatives: think about the likelihood function and how you would maximize it).

We want to maximize

$$L(x_1, x_2, \ldots x_n \mid a, b, c, d) = \prod_{i=1}^{n} \frac{1}{(b-a)+(d-c)} = \left(\frac{1}{(b-a)+(d-c)}\right)^n$$
$$where\ \forall i, x_i \in [a,b] \cup [c,d], 0\ otherwise$$

We need $(b-a)+(d-c)$ to be as small as possible while having all $x_i \in [a,b] \cup [c,d]$ to maximize, so we need $[a,b]$ and $[c,d]$ to be as small as possible.

Sort the x_i such that $x_1 \leq x_2 \leq \cdots \leq x_n$.

$$\hat{a} = x_{min} = x_1$$
$$\hat{d} = x_{max} = x_n$$
$$(\hat{b}, \hat{c}) = (x_i, x_{i+1})\ where\ i = \underset{2 \leq i \leq n-2}{\operatorname{argmax}}(x_{i+1} - x_i)$$

(In other words, the (x_i, x_{i+1}) that are the farthest apart when sorted).

3. Let's say a random variable $X \sim Exp2(\mu), \mu > 0$ if it has the following density:

$$f_X(x) = \begin{cases} \frac{1}{\mu} e^{-x/\mu}, & x \geq 0 \\ 0, & otherwise \end{cases}$$

You are given that $E[X] = \mu$ and $Var(X) = \mu^2$.

a) Suppose we have iid samples, $x_1, \ldots, x_n \sim Exp2(\mu)$. Show that the maximum likelihood estimator for μ is $\widehat{\mu_n} = \frac{1}{n}\sum_{i=1}^{n} x_i$, . You need not verify that your estimator is a maximum.

$$L(x_1, x_2, \ldots x_n \mid \mu) = \prod_{i=1}^{n} \frac{1}{\mu} e^{-x_i/\mu}$$

$$\ln(L(x_1, x_2, \ldots x_n \mid \mu)) = \sum_{i=1}^{n} \left(-\ln(\mu) - \frac{x_i}{\mu}\right)$$

$$\frac{\partial}{\partial \mu}\left[\ln(L(x_1, x_2, \ldots x_n \mid \mu))\right] = \sum_{i=1}^{n} \left(-\frac{1}{\mu} + \frac{x_i}{\mu^2}\right) = 0$$

$$\sum_{i=1}^{n} \left(-1 + \frac{x_i}{\widehat{\mu_n}}\right) = -n + \frac{1}{\widehat{\mu_n}} \sum_{i=1}^{n} x_i = 0 \rightarrow \widehat{\mu_n} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

b) Let's consider our estimator from part a). Is your maximum likelihood estimator unbiased? Prove it.

It is unbiased.

$$E[\widehat{\mu_n}] = E\left[\frac{1}{n}\sum_{i=1}^{n} X_i\right] = \frac{1}{n}\sum_{i=1}^{n} E[X_i] = \frac{1}{n} n\mu = \mu$$

c) MLE's have a lot of nice properties, out of the scope of this class. Unfortunately, unbiasedness is not one of them. Give an example of a distribution and a parameter for that distribution where the maximum likelihood estimator is biased. (Hint: You've seen at least two examples in class or on homework).

Examples may vary. Two possibilities are: either the estimator for σ^2 in the normal distribution, $N(\mu, \sigma^2)$, or the estimator for θ in the one-parameter uniform pdf, $Unif(0, \theta)$.

d) Find $Var(\widehat{\mu_n})$.

$$Var(\widehat{\mu_n}) = Var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{1}{n^2} n\mu^2 = \frac{\mu^2}{n}$$

e) The mean squared error of an estimator is defined as $MSE(\hat{\theta}) = Var(\hat{\theta}) + Bias^2(\hat{\theta})$. It measures how much an estimator deviates from the true parameter, whereas the variance of the estimator just measures how much the estimator deviates from its own expectation. The bias of an estimator is determined by $Bias(\hat{\theta}) = E[\hat{\theta}] - \theta$. Find $MSE(\widehat{\mu_n})$.

Since our estimator was unbiased, $E[\widehat{\mu_n}] - \mu = 0$, so $MSE(\hat{\theta}) = Var(\hat{\theta}) = \frac{\mu^2}{n}$.

f) We say an estimator $\widehat{\theta_n}$ is consistent for θ if it converges in probability to θ. That is,

$$\forall\varepsilon > 0, \lim_{n\to\infty} P\left(\left|\widehat{\theta_n} - \theta\right| \geq \varepsilon\right) = 0$$

In English, this means that our estimator approaches the true value θ as we take more and more samples. Although MLE's may be biased, MLE's are guaranteed to be consistent (which implies asymptotic unbiasedness, or "unbiased in the limit"). Use Chebyshev's Inequality to show that the MLE for $Exp2(\mu)$, $\widehat{\mu_n}$, is consistent for μ. We have started the proof below.

<u>Proof:</u>

Fix $\varepsilon > 0$. By Chebyshev's inequality, $0 \leq P(|\widehat{\mu_n} - \mu| \geq \varepsilon) \leq \frac{Var(\widehat{\mu_n})}{\varepsilon^2} = \frac{\mu^2}{n\varepsilon^2}$. As $n \to \infty$, $\frac{\mu^2}{n\varepsilon^2} \to 0$, so we have $0 \leq P(|\widehat{\mu_n} - \mu| \geq \varepsilon) \leq 0$, and by the Sandwich Theorem, $\lim_{n\to\infty} P(|\widehat{\mu_n} - \mu| \geq \varepsilon) = 0$. So our estimator $\widehat{\mu_n}$ is consistent.

Q.E.D.

4.

a) Suppose we have the inequality, $a_1 + a_2 + \cdots + a_n \leq k$, where $k \geq n$. What is the number of solutions to this inequality, with the constraint that each a_i must be a positive integer and that k is also an integer? (Hint 1: Modify the inequality to get an equivalent one in which each a_i is constrained to be nonnegative instead of positive). (Hint 2: introduce a new term a_{n+1} which will take the "remainder").

This is equivalent to $a_1 + a_2 + \cdots + a_n \leq k - n$, where $a_i \geq 0$ instead of $a_i \geq 1$ (check that this is equivalent). Then, introduce a_{n+1} to take the remainder to get $a_1 + a_2 + \cdots + a_n + a_{n+1} = k - n$ (check that this is also equivalent). Now we can do exactly what we did in homework 1, using a technique called "stars and bars", to get the number of solutions:

$$\binom{(k-n) + (n+1) - 1}{(n+1) - 1} = \binom{k}{n}$$

This is because we have $k - n$ "stars" and n "bars" to divide the "stars". So there are a total of k "stars" and "bars", and we choose the location of the "bars" (an equivalent expression would have chosen the location of the "stars").

b) Suppose we have the same setup as in part a), but we flipped the direction of our constraint inequality to be $k < n$. How many solutions are there now with this reversed constraint?

There are none. Their sum must be at least n, since each $a_i \geq 1$.

5. Every week, $20{,}000$ students roll a $10{,}000$-sided fair die, numbered 1 to $10{,}000$, to see if they can get their GPA changed to a 4.0. If they roll a 1, they win (they get their GPA changed). You may assume each student's roll is independent. Let X be the number of students who win.

a) For any given week, give the appropriate probability distribution (including parameter(s)), and find the expected number of students who win.

$$X \sim Bin\left(n = 20000, p = \frac{1}{10000}\right) \rightarrow E[X] = np = 20000\left(\frac{1}{10000}\right) = 2$$

b) For any given week, find the exact probability that at least 2 students win. Give your answer to 5 decimal places.

$$P(X \geq 2) = 1 - P(X = 0) - P(X = 1) = 1 - \binom{n}{0} p^0 (1-p)^n - \binom{n}{1} p^1 (1-p)^{n-1}$$
$$\approx 1 - 0.40599 \approx 0.59401$$

c) For any given week, estimate the probability that at least 2 students win, using the Poisson approximation. Give your answer to 5 decimal places.

$$X' \sim Poi(np) \sim Poi(2)$$

$$P(X \geq 2) \approx P(X' \geq 2) = 1 - P(X' = 0) - P(X' = 1) = 1 - \frac{e^{-2}2^0}{0!} - \frac{e^{-2}2^1}{1!} \approx 0.59399$$

d) For any given week, estimate the probability that at least 2 students win, using the Normal approximation. Give your answer to 4 decimal places.

$$X'' \sim N\big(np, np(1-p)\big) \sim N\left(2, \frac{19998}{10000}\right)$$

We'll say this is essentially $N(2,2)$. Use the continuity correction.

$$P(X \geq 2) \approx P(X'' \geq 1.5) = P\left(Z \geq \frac{1.5-2}{\sqrt{2}}\right) = P(Z \geq -0.3535534)$$
$$= \phi(0.3535534) \approx 0.6406 \; or \; 0.6368 \; or \; 0.6382$$

6. Suppose X has the following probability mass function:

$$p_X(x) = \begin{cases} c, & x = 0 \\ 2c, & x = \frac{\pi}{2} \\ c, & x = \pi \\ 0, & otherwise \end{cases}$$

Trig Reference

Angle	Sine	Cosine
0	0	1
$\pi/2$	1	0
π	0	-1

a) Suppose $Y_1 = \sin(X)$. Find $E[Y_1^2]$.

Probabilities must sum to 1, so $c = 1/4$.

$$E[Y_1^2] = E[\sin^2(X)] = \frac{1}{4} * \sin^2(0) + \frac{1}{2} * \sin^2\left(\frac{\pi}{2}\right) + \frac{1}{4} * \sin^2(\pi) = \frac{1}{2}$$

b) Suppose $Y_2 = \cos(X)$. Find $E[Y_2^2]$.

$$E[Y_2^2] = E[\cos^2(X)] = \frac{1}{4} * \cos^2(0) + \frac{1}{2} * \cos^2\left(\frac{\pi}{2}\right) + \frac{1}{4} * \cos^2(\pi) = \frac{1}{2}$$

c) Suppose $Y = Y_1^2 + Y_2^2 = \sin^2(X) + \cos^2(X)$. Before any calculation, what do you think $E[Y]$ should be? Find $E[Y]$, and see if your hypothesis was correct. (Recall for any real number x, $\sin^2(x) + \cos^2(x) = 1$).

I expect the answer to be 1, since for any real number x, $\sin^2(x) + \cos^2(x) = 1$, but I'm not sure since these are random variables and not real numbers.

$$E[Y] = E[Y_1^2 + Y_2^2] = E[Y_1^2] + E[Y_2^2] = \frac{1}{2} + \frac{1}{2} = 1$$

d) Let W be any discrete random variable with probability mass function $p_W(w)$. Then, $E[\sin^2(W) + \cos^2(W)] = 1$. Is this statement always true? If so, prove it. If not, give a counterexample by giving a probability mass function for a discrete random variable W for which the statement is false.

This is true. Recall for a discrete random variable, $E[g(X)] = \sum_x g(x)p_X(x)$.

$$E[\sin^2(W) + \cos^2(W)] = \sum_w (\sin^2(w) + \cos^2(w))p_W(w) = \sum_w 1 * p_W(w) = \sum_w p_W(w) = 1$$

7. Suppose any given CSE 312 student is three times more likely to come to class regularly than not. A student who doesn't come to class regularly is twice as likely to have failuritis as a student who does.

a) What is the probability a student doesn't come to class regularly, given that they have failuritis?

Let $A = \text{"}student\ doesn't\ come\ to\ class\ regularly\text{"}$
Let $B = \text{"}student\ has\ failuritis\text{"}$

Then, $3P(A) = P(A^C) \rightarrow P(A) = \frac{1}{4}$.

We are also given $P(B|A) = 2P(B|A^C)$. This implies that $\frac{P(B \cap A)}{P(A)} = 2\frac{P(B \cap A^C)}{P(A^C)}$.
Then, $\frac{3}{4}P(B \cap A) = 2 \cdot \frac{1}{4}P(B \cap A^C) \rightarrow \frac{3}{2}P(B \cap A) = P(B \cap A^C)$.

$P(B) = P(B \cap A) + P(B \cap A^C) = P(B \cap A) + \frac{3}{2}P(B \cap A) = \frac{5}{2}P(B \cap A)$.

Therefore, $P(A|B) = \frac{P(B \cap A)}{P(B)} = \frac{P(B \cap A)}{\frac{5}{2}P(B \cap A)} = \frac{2}{5}$.

b) What is the probability a student comes to class regularly, given that they have failuritis?

$P(A^C|B) = 1 - P(A|B) = 1 - \frac{2}{5} = \frac{3}{5}$.

8.

a) Let $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$, where X and Y are independent. Show that their sum $Z = X + Y$ is $Poi(\lambda_1 + \lambda_2)$. (Hint 1: Start with the pmf for Z, $P(Z = n)$). (Hint 2: Use the binomial theorem).

$$P(X+Y=n) = \sum_{k=0}^{n} P(X=k \cap Y=n-k) = \sum_{k=0}^{n} P(X=k)P(Y=n-k)$$
$$= \sum_{k=0}^{n} \frac{e^{-\lambda_1}\lambda_1^k}{k!}\frac{e^{-\lambda_2}\lambda_2^{n-k}}{(n-k)!} = \frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n! * \lambda_1^k\lambda_2^{n-k}}{k!\,(n-k)!}$$

Recall from the binomial theorem that

$$(\lambda_1+\lambda_2)^n = \sum_{k=0}^{n}\frac{n! * \lambda_1^k\lambda_2^{n-k}}{k!\,(n-k)!}$$

So

$$P(X+Y=n) = \frac{e^{-(\lambda_1+\lambda_2)}}{n!}(\lambda_1+\lambda_2)^n$$

This is the probability mass function for $Poi(\lambda_1 + \lambda_2)$.

b) Suppose $X_1, X_2, \dots, X_n$ are iid $\sim Poi(\lambda)$. Their sum $X = X_1 + \cdots + X_n$ is Poisson with what rate parameter λ'?

The previous argument can be extended by induction to show that $\lambda' = n\lambda$.

c) Let $X \sim Poi(\lambda_1)$ and $Y \sim Poi(\lambda_2)$, where X and Y are independent. In part a), we showed that $Z = X + Y$ is $Poi(\lambda_1 + \lambda_2)$. Prove that $P(X = k|Z = n)$ is the probability mass function for a binomial random variable, and specify its parameters. You may use any results from previous parts. We have started the proof below. (Hint: If $X = k$ and $Z = n$, what does that say about Y?).

$$P(X=k|Z=n) = \frac{P(X=k\cap Z=n)}{P(Z=n)}$$
$$= \frac{P(X=k\cap Y=n-k)}{P(Z=n)}$$
$$= \frac{\frac{e^{-\lambda_1}\lambda_1^k}{k!}\frac{e^{-\lambda_2}{\lambda_2}^{n-k}}{(n-k)!}}{\frac{e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n}{n!}}$$
$$= \binom{n}{k}\frac{\lambda_1^k{\lambda_2}^{n-k}}{(\lambda_1+\lambda_2)^n}$$
$$= \binom{n}{k}\left(\frac{\lambda_1}{\lambda_1+\lambda_2}\right)^k\left(\frac{\lambda_2}{\lambda_1+\lambda_2}\right)^{n-k}$$

This is the probability mass function for $Bin\left(n, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$.

9. Recall that for a continuous random variable Y, we have that

$$E[Y] = \int_{-\infty}^{\infty} y f_Y(y) dy$$

Suppose Y is a non-negative continuous random variable. Show that $E[Y] = \int_0^{\infty} (1 - F_Y(y)) dy$, where $F'_Y(y) = f_Y(y)$. (Hint: use double integrals).

$$\int_0^{\infty} (1 - F_Y(y)) dy = \int_0^{\infty} P(Y > y) dy \;\; [by\ definition]$$

$$= \int_0^{\infty} \int_y^{\infty} f_Y(x) dx\, dy \;\; \left[since\ P(Y > y) = \int_y^{\infty} f_Y(x) dx\right]$$

$$= \int_0^{\infty} \int_0^{x} f_Y(x) dy\, dx \;\; [switching\ order\ of\ integration]$$

$$= \int_0^{\infty} f_Y(x) \left(\int_0^{x} dy \right) dx \;\; [since\ f_Y(x)\ is\ constant\ with\ respect\ to\ y]$$

$$= \int_0^{\infty} f_Y(x) x dx = E[Y]$$

10. Let P be a continuous uniform random variable distributed on $[a, b]$, where $0 < a < b < 1$. Let $N \sim Poi(\lambda)$. Suppose $X_1, \ldots, X_N$ are iid $Ber(P)$. So we have N iid Bernoulli random variables, each with success parameter P, where both N and P are random variables themselves! Let

$$X = \sum_{i=1}^{N} X_i$$

Find $E[X]$ in terms of a, b, and λ. (Hint: Use the Law of Total Expectation).

$$E[X] = \sum_{n=0}^{\infty} E[X \mid N = n] P(N = n) \;\; [by\ law\ of\ total\ expectation]$$

$$= \sum_{n=0}^{\infty} E\left[\sum_{i=1}^{N} X_i \mid N = n\right] P(N = n)$$

$$= \sum_{n=0}^{\infty} E\left[\sum_{i=1}^{n} X_i\right] P(N = n)$$

$$= \sum_{n=0}^{\infty} n E[X_1] P(N = n)$$

$$= E[X_1] \sum_{n=0}^{\infty} n P(N = n) = E[X_1] E[N] = \frac{a + b}{2} \lambda$$