CSE 312: Foundations of Computing II
Quiz Section #10: Review Questions for Final Exam (solutions)

1. (**Confidence Intervals, CLT**) Let $X_1, ..., X_n$ be iid with unknown mean θ and known variance σ^2. Assume n is sufficiently large, and our maximum likelihood estimate for θ is the sample mean $\hat{\theta} = \bar{X}_n = \frac{1}{n}\sum_{i=1}^n X_i$. Our estimator is wrong with probability 1; that is, $\mathbb{P}(\hat{\theta} = \theta) = 0$ since θ is a real number. Find z such that $\mathbb{P}\left(\theta \in \left[\hat{\theta} - \frac{z\sigma}{\sqrt{n}}, \hat{\theta} + \frac{z\sigma}{\sqrt{n}}\right]\right) \geq 1 - \alpha$, for any $\alpha \in (0, 1)$. We call $\left[\hat{\theta} - \frac{z\sigma}{\sqrt{n}}, \hat{\theta} + \frac{z\sigma}{\sqrt{n}}\right]$ a $100(1-\alpha)\%$ confidence interval for θ.

   First notice that if $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1) \equiv Z$. But $Z \equiv -Z$, since it is symmetric. So below, we will be using $-Z$ in place of Z, but they have the same distribution.

   By the CLT, $\hat{\theta} \approx \mathcal{N}(\theta, \frac{\sigma^2}{n})$. We have

$$\mathbb{P}\left(\theta \in \left[\hat{\theta} - \frac{z\sigma}{\sqrt{n}}, \hat{\theta} + \frac{z\sigma}{\sqrt{n}}\right]\right) = \mathbb{P}\left(\hat{\theta} - \frac{z\sigma}{\sqrt{n}} \leq \theta \leq \hat{\theta} + \frac{z\sigma}{\sqrt{n}}\right)$$

   Standardizing (subtract $\hat{\theta}$ and divide by $\frac{\sigma}{\sqrt{n}}$ on all sides) gives

$$\mathbb{P}(-z \leq Z \leq z) = \Phi(z) - \Phi(-z) = \Phi(z) - (1 - \Phi(z)) = 2\Phi(z) - 1$$

   We want this $\geq 1 - \alpha$, and solving for z gives $\Phi^{-1}(1 - \frac{\alpha}{2})$.

2. (**Transformation**) Let $X \sim Unif(-1, 1)$, and $Y = X^2$. What is $f_Y(y)$?

   First note that $\Omega_Y = [0, 1]$. If $y \in [0, 1]$, we have

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(X^2 \leq y) = \mathbb{P}(-\sqrt{y} \leq X \leq \sqrt{y}) = \frac{2\sqrt{y}}{2} = \sqrt{y}$$

   Hence,

$$f_Y(y) = \begin{cases} \frac{1}{2\sqrt{y}} & y \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

3. (**Convolution**) Suppose $Z = X + Y$, where $X \perp Y$. Z is called the convolution of two random variables. If X, Y, Z are discrete,

$$p_Z(z) = \mathbb{P}(X + Y = z) = \sum_x \mathbb{P}(X = x \cap Y = z - x) = \sum_x p_X(x)\, p_Y(z - x)$$

   If X, Y, Z are continuous,

$$F_Z(z) = \mathbb{P}(X + Y \leq z) = \int_{-\infty}^{\infty} \mathbb{P}(Y \leq z - X \mid X = x) f_X(x) dx = \int_{-\infty}^{\infty} F_Y(z - x) f_X(x) dx$$

   Suppose $X_1 \sim \mathcal{N}(\mu_1, \sigma_1^2)$ and $X_2 \sim \mathcal{N}(\mu_2, \sigma_2^2)$.

(a) Find an expression for $\mathbb{P}(X_1 < 2X_2)$ using a similar idea to convolution, in terms of $F_{X_1}, F_{X_2}, f_{X_1}, f_{X_2}$. (Your answer will be in the form of a single integral, and requires no calculations – do not evaluate it).

We use the continuous version of the "Law of Total Probability" to integrate over all possible values of X_2. Take the probability that $X_1 < 2X_2$ given that value of X_2, times the density of X_2 at that value.

$$\mathbb{P}(X_1 < 2X_2) = \int_{-\infty}^{\infty} \mathbb{P}(X_1 < 2X_2 \mid X_2 = x_2) f_{X_2}(x_2)\, dx_2 = \int_{-\infty}^{\infty} F_{X_1}(2x_2)\, f_{X_2}(x_2)\, dx_2$$

(b) Find s, where $\Phi(s) = \mathbb{P}(X_1 < 2X_2)$ using the "reproductive" property of normal distributions.

Let $X_3 = X_1 - 2X_2$, so that $X_3 \sim \mathcal{N}(\mu_1 - 2\mu_2, \sigma_1^2 + 4\sigma_2^2)$ (by the reproductive property of normal distributions)

$$\mathbb{P}(X_1 < 2X_2) = \mathbb{P}(X_1 - 2X_2 < 0) = \mathbb{P}(X_3 < 0) = \mathbb{P}\left(\frac{X_3 - (\mu_1 - 2\mu_2)}{\sqrt{\sigma_1^2 + 4\sigma_2^2}} < \frac{0 - (\mu_1 - 2\mu_2)}{\sqrt{\sigma_1^2 + 4\sigma_2^2}} \right)$$

$$= \mathbb{P}\left(Z < \frac{2\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + 4\sigma_2^2}} \right) = \Phi\left(\frac{2\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + 4\sigma_2^2}} \right) \rightarrow s = \frac{2\mu_2 - \mu_1}{\sqrt{\sigma_1^2 + 4\sigma_2^2}}$$

4. (**Counting, Discrete RVs**) Suppose we have N items in a bag, K of which are successes. Suppose we draw (without replacement) until we have k successes, $k \leq K \leq N$. Let X be the number of draws until the k^{th} success. What is the range Ω_X? What is $p_X(n) = P(X = n)$? (We say X is a "negative hypergeometric" random variable).

First, let's think about the possible values of X. The smallest possible value of X is k, since we need at least k draws to get k successes. The largest possible value of X is $(N - K) + k$, since in the worst case, we could draw all $(N - K)$ failures before drawing k successes. Thus X can take on these values:

$$\Omega_X = \{k, k+1, \dots N - K + k\}$$

Now, what does it mean for $X = n$? This means that out of the first $n - 1$ draws, there must have been exactly $k - 1$ successes, with the remaining $(n - 1) - (k - 1) = n - k$ draws being failures. There are $\binom{K}{k-1}$ ways to select a set of $k - 1$ successes, $\binom{N-K}{n-k}$ ways to select a set of $n - k$ failures, and there are a total of $\binom{N}{n-1}$ ways to select $n - 1$ items overall.

Finally, the n-th draw must have been a success. By this point, there are $N - (n - 1)$ total items remaining in the bag, and $K - (k - 1)$ successes remaining. Putting this all together, we have

$$p_X(n) = \mathbb{P}(X = n) = \frac{\binom{K}{k-1}\binom{N-K}{n-k}}{\binom{N}{n-1}} \frac{K - (k-1)}{N - (n-1)}, \quad n = k, k+1, \dots, N - K + k$$

5. (**Bias**) Suppose $x_1, x_2, \dots, x_n$ are independent, identically distributed samples from the continuous distribution $\text{Unif}(0, \theta)$. Consider the estimator $\hat{\theta} = \frac{3}{n}\sum_{i=1}^n x_i$ of θ. Is $\hat{\theta}$ unbiased? If not, find a constant c such that $c\hat{\theta}$ is unbiased and prove that it is unbiased.

$$\mathbb{E}[\hat{\theta}] = \mathbb{E}\left[\frac{3}{n}\sum_{i=1}^n x_i\right] = \frac{3}{n}\sum_{i=1}^n \mathbb{E}[x_i] = \frac{3}{n}\cdot n \cdot \frac{\theta}{2} = \frac{3}{2}\cdot\theta$$

so $\hat{\theta}$ is biased. Let $c = \frac{2}{3}$. Then

$$\mathbb{E}[c\hat{\theta}] = \mathbb{E}\left[\frac{2}{3}\cdot\hat{\theta}\right] = \frac{2}{3}\mathbb{E}[\hat{\theta}] = \frac{2}{3}\cdot\frac{3}{2}\cdot\theta = \theta$$

so $\hat{\theta}' = c\hat{\theta} = \frac{2}{n}\sum_{i=1}^n x_i$ is an unbiased estimator. Note, though, that it may be a very poor estimator for θ, since some of the samples could even be greater than $\hat{\theta}'$: try it out on the samples $1, 2, 9$. An estimator may not be good just because it is unbiased. Compare this to the MLE you got on your homework.

6. (**Expectation of a Function**) You flip a fair coin independently and count the number of flips until the first tail, including that tail flip in the count. If the count is n, you receive 2^n dollars. What is the expected amount you will receive? How much would you be willing to pay at the start to play this game?

The expected amount is ∞. Let N be the number of flips until the first tail, so $N \sim Geo(1/2)$, and $p_N(n) = 1/2^n$ for $n \in \mathbb{N}$. Hence $\mathbb{E}[2^N] = \sum_{n=1}^\infty 2^n \dfrac{1}{2^n} = \sum_{n=1}^\infty 1 = \infty$. In theory, you should be willing to pay any finite amount of money to play this game, but I admit I would be nervous to pay a lot. For instance, if you pay \$1000, you will lose money unless the first 9 flips are all heads. With high probability you will lose money, and with low probability you will win a lot of money.

7. (**Law of Total Probability, Counting**) Suppose A and B are random, independent, nonempty subsets of $\{1, 2, \dots, n\}$, where each nonempty subset is equally likely to be chosen as A or B. What is $\mathbb{P}(\max(A) = \max(B))$?

We know $\mathbb{P}[\max(A) = k] = \frac{2^{k-1}}{2^n - 1}$, since for the max to be k, A needs to contain k and no element greater than k, and there are 2^{k-1} subsets of $\{1, 2, \dots, k-1\}$. Now use the law of total probability:

$$\begin{aligned}
\mathbb{P}[\max A = \max B] &= \sum_{k=1}^n \mathbb{P}[\max A = \max B \mid \max B = k]\cdot\mathbb{P}[\max B = k] \\
&= \sum_{k=1}^n \mathbb{P}[\max A = k]\cdot\mathbb{P}[\max B = k] \\
&= \sum_{k=1}^n \left(\frac{2^{k-1}}{2^n - 1}\right)^2 = \frac{1}{(2^n-1)^2}\sum_{k=1}^n 4^{k-1} = \frac{4^n - 1}{3(2^n-1)^2} = \frac{2^n+1}{3(2^n-1)}
\end{aligned}$$

8. (**Linearity of Expectation**) Suppose A and B are random, independent (possibly empty) subsets of $\{1, 2, \dots, n\}$, where each subset is equally likely to be chosen as A or B. Consider $A\Delta B = (A \cap B^C) \cup (B \cap A^C) = (A \cup B) \cap (A^C \cup B^C)$, i.e., the set containing elements that are in exactly one of A and B. Let X be the random variable that is the size of $A\Delta B$. What is $\mathbb{E}[X]$?

For $i = 1, 2, \ldots, n$, let X_i be the indicator of whether $i \in A \Delta B$. Then $X_i \sim \text{Ber}(\frac{1}{2})$, and $X = \sum_{i=1}^n X_i$, so $\mathbb{E}[X] = \mathbb{E}[\sum_{i=1}^n X_i] = \frac{n}{2}$.

9. (**Transformation**) Let X be a continuous random variable with invertible CDF F_X, and let $Y = F_X(X)$. Show that $Y \sim Unif(0, 1)$.

   First note that $\Omega_Y = [0, 1]$. If $y \in [0, 1]$, we have

$$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(F_X(X) \leq y) = \mathbb{P}(X \leq F_X^{-1}(y)) = F_X(F_X^{-1}(y)) = y$$

   Therefore

$$f_Y(y) = \begin{cases} 1 & y \in [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

   and so $Y \sim Unif(0, 1)$.

10. A European city's temperature is modeled as a random variable with mean μ and standard deviation σ, measured on the Celsius scale. A day is described as "ordinary" if the temperature during that day remains within one standard deviation of the mean.

    (a) Give formulas for the mean and variance, if temperature is measured on the Fahrenheit scale. The formula for conversion is F = 32 + 1.8C.

$$\mathbb{E}[F] = 32 + 1.8\mu, Var(F) = (1.8)^2\sigma^2$$

    (b) From your formulas in part (a), give formulas for the temperature range for an ordinary day on the Fahrenheit scale.

$$32 + 1.8\mu - 1.8\sigma \text{ to } 32 + 1.8\mu + 1.8\sigma$$

11. During each day, the probability that your computer's operating system crashes at least once is 5%, independent of every other day. You are interested in the probability of at least 45 crash-free days out of the next 50 days.

    If X is the number of days with a crash, then $X \sim \text{Bin}(50, 0.05)$, with $\mu = 2.5$ and $\sigma^2 = 2.375$.

    (a) Find the probability of interest by using the normal approximation to the binomial.

    $\mathbb{P}(X \leq 5) = \mathbb{P}(X < 5.5) = \mathbb{P}\left(\frac{X-2.5}{\sqrt{2.375}} < \frac{5.5-2.5}{\sqrt{2.375}}\right) \approx \Phi\left(\frac{3}{\sqrt{2.375}}\right) \approx 0.9744$. If you got 0.9474, you forgot the continuity correction.

    (b) Find the probability of interest by using the Poisson approximation to the binomial.

    X is approximately Poi(2.5), so $\mathbb{P}(X \leq 5) \approx \sum_{k=0}^{5} 2.5^k \cdot e^{-2.5}/k! \approx 0.9580$. But if you were going to sum 6 such terms, you might as well get a more accurate answer and sum the correct binomial terms instead, giving the answer 0.9622.

12. Consider the line segment $[0, L]$. Let $X \sim \text{Exp}(4/L)$. If $0 \leq X \leq L$, the line segment $[0, L]$ is split into two at the point X (yielding one piece of length X and one piece of length $L - X$), otherwise it is split into two at the point L (yielding one piece of length L and one piece of length 0). Give your answers to 3 significant digits.

    (a) Find the probability that the ratio of the shorter to the longer segment is less than 1/3.

    $$\mathbb{P}(X < L/4) + \mathbb{P}(X > 3L/4) = (1 - e^{-1}) + e^{-3} \approx 0.632 + 0.050 = 0.682$$

    (b) What is the probability that X is less than 0 or greater than L?

    $$\mathbb{P}(X > L) = e^{-4} \approx 0.018$$

13. A computer network consisting of n computers is to be formed by connecting each computer to each of the others by a direct ("point-to-point") network cable.

    (a) How many network cables are needed?

    $$\binom{n}{2}$$

    (b) Unfortunately, some of the cables may be faulty ("dead") while others are OK ("alive"). How many different "connectivity patterns" are possible? (E.g., "the cable between computers 1 and 3 is alive, but no others are" is one pattern; "between 1 and 4, but no others" is a different pattern; "only the cable between 1 and 4 is dead" is a third pattern, etc.)

    $$2^{\binom{n}{2}}$$

    (c) Assuming that there is at least one "live" cable connected to every computer, show that there are at least two computers in the network that are directly connected to the same number of other computers via live cables.

    Use the pigeonhole principle.

14. Alice, Bob, and Carol repeatedly take turns rolling a fair die. Alice begins, Bob always follows Alice, Carol always follows Bob, and Alice always follows Carol. Find the probability that Carol will be the first one to roll a six.

    $$\left(\frac{5}{6}\right)^2 \left(\frac{1}{6}\right) + \left(\frac{5}{6}\right)^5 \left(\frac{1}{6}\right) + \ldots \approx 0.275$$

15. Consider the line segment $[0, L]$. Let $X \sim \mathrm{N}(L/2, L^2/16)$. If $0 \leq X \leq L$, the line segment $[0, L]$ is split into two at the point X (yielding one piece of length X and one piece of length $L - X$), otherwise it is split into two at the point L (yielding one piece of length L and one piece of length 0). Give your answers to 3 significant digits.

    (a) Find the probability that the ratio of the shorter to the longer segment is less than 1/3.

    0.3174

    (b) What is the probability that X is less than 0 or greater than L?

    0.0456

16. The number of seconds a server takes to finish a job is modeled as a random variable X from an unknown distribution. You would like to be able to guarantee clients that, with high probability, jobs will be finished within 50 seconds. What is the best guarantee you could give if:

    (a) You assume that X has mean 25.

    $\mathbb{P}(X \geq 50) \leq 25/50 = 1/2$ by Markov's inequality, so $\mathbb{P}(X < 50) \geq 1 - 1/2 = 1/2$.

    (b) You assume that X has mean 25 and variance 25.

    $\mathbb{P}(X \geq 50) = \mathbb{P}(X - 25 \geq 25) \leq 25/(25 + 25^2) = 1/26$ by Cantelli's inequality, so $\mathbb{P}(X < 50) \geq 25/26$.

    (c) You assume that $X \sim \mathrm{Poi}(25)$. (Hint: use the Normal approximation of the Poisson. Why is it reasonable to approximate Poi(25) by a normal distribution? It follows from the Central Limit Theorem, since it turns out that a Poisson random variable with $\lambda = 25$ is the sum of 25 independent Poisson random variables each with $\lambda = 1$. See `https://onlinecourses.science.psu.edu/stat414/node/180`.)

    Approximate X by $Y \sim N(25, 25)$. $\mathbb{P}(Y \leq 50) = \mathbb{P}(Y < 50.5) = \mathbb{P}((Y - 25)/5 < 5.1) = \Phi(5.1) \approx 1$. Why is there a continuity correction? It's because the Poisson is a discrete distribution and we're approximating it by a continuous distribution.

17. A frog starts at position 0 on a line and at each second t jumps X_t cm, where the X_t are all i.i.d. according to the following probability mass function:

$$\begin{aligned} p_X(-2) &= 1/6 \\ p_X(-1) &= 1/3 \\ p_X(1) &= 1/6 \\ p_X(2) &= 1/3 \end{aligned}$$

    Use the central limit theorem to estimate the probability that, after 100 jumps, the frog is at a negative position.

    $\mathbb{E}[X_t] = 1/6$ and $\mathrm{Var}(X_t) = 89/36$. If $X = \sum_{t=1}^{100} X_t$, $\mathbb{E}[X] \approx 16.67$ and $\mathrm{Var}(X) \approx 247$. So $\mathbb{P}(X < 0) \approx \Phi(-1.09) \approx 0.1379$. If you got $\Phi(-1.06)$, you forgot the continuity correction.

18. Chebyshev's inequality implies that the proportion of observations that are at most 3 standard deviations from the mean is at least p. Determine the value of p.

8/9

19. You throw a dart at a circular target of radius $r = 5$ inches. Your aim is such that the dart is equally likely to hit any point in the target. For each throw, you win \$1 if the dart strikes within 2 inches of the target's center. Let W be your total winnings for 100 independent throws. Use the Chernoff bound to get an upper bound on the probability that you win at least \$24. (The Chernoff bound is sometimes given in the form $\mathbb{P}(X > (1+\delta)\mu) \leq \ldots$, but the same bound actually also holds in the form $\mathbb{P}(X \geq (1+\delta)\mu) \leq \ldots$.)

$W \sim \text{Bin}(100, 4/25)$, with $\mu = 16$. $\mathbb{P}(W \geq 24) = \mathbb{P}(W \geq 1.5\mu) \leq e^{-16/12} < 0.264$.

20. Suppose $x_1, x_2, \ldots, x_n$ are independent samples from $\text{Bin}(N, p)$, where the parameter N is known to you but p is unknown.

    (a) What is the maximum likelihood estimator for p? Don't forget to prove that it is a maximum of the likelihood function.

$$\hat{p} = \frac{1}{nN} \sum_{i=1}^{n} x_i$$

    (b) Is your answer to part (a) a biased or unbiased estimator?

    Unbiased

21. For any individual x born in Transylvania with a vampire father, there is a 50% chance that x is a vampire, independently for each birth. These are the only conditions under which a new vampire can be created. 75% of the Transylvanian males are vampires. Suppose Igor, a man who has lived in Transylvania his whole life, has three children that are not vampires.

    (a) What is the probability that Igor is a vampire?

3/11

    (b) If Igor has a fourth child, what is the probability that child will be a vampire?

3/22

22. A bridge deck consists of 52 cards divided into 4 suits of 13 ranks each. A bridge hand consists of 13 cards from a bridge deck. Suppose that the bridge cards are well shuffled and dealt. What is the probability that your bridge hand is already sorted when you pick it up, given that you have been dealt at least two cards in each of the 4 suits? By "sorted" I mean that the cards of any one suit are adjacent to each other, and the cards of each

suit are sorted by rank, with ascending ranks either from left to right or from right to left in your hand. The 4 suits can be in any order in your hand, and different suits can sorted in different directions.

$$\frac{4! \cdot 2^4}{13!} \approx 6 \times 10^{-8}$$

23. Let X be a continuous random variable with probability density function

$$f(x) = \begin{cases} 2x & , \text{ if } 0 \le x \le 1 \\ 0 & , \text{ otherwise} \end{cases} .$$

(a) Find $\mathbb{E}\left[\frac{1}{X}\right]$.

$$\mathbb{E}\left[\frac{1}{X}\right] = \int_0^1 \frac{1}{x} \cdot 2x dx = 2$$

(b) Compute $\mathbb{P}(X = 0.5)$.

0

24. Bob is teaching Alice how to play his new favorite game. In each round, Bob shoots an arrow at the tires of Alice's car. He hits with probability p, independent of previous rounds. If he hits a tire, he gets 10 points. If he misses, he loses 5. Let X be Bob's score after n rounds.

(a) What is $\mathbb{E}[X]$?

Let X_i be the number of points Bob gets for the i-th round.

$$\begin{aligned} \mathbb{E}[X_i] &= 10p + (-5)(1-p) = 15p - 5 \\ \mathbb{E}[X] &= E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i] = 15np - 5n \end{aligned}$$

(b) What is $\mathrm{Var}(X)$?

$$\begin{aligned} Var(X_i) &= \mathbb{E}[X_i^2] - (\mathbb{E}[X_i])^2 = (10^2 p + (-5)^2(1-p)) - (15p-5)^2 \\ &= (100p + 25 - 25p) - (225p^2 - 150p + 25) = 225p(1-p) \\ Var(X) &= Var\left(\sum_{i=1}^{n} X_i\right) = \sum_{i=1}^{n} Var(X_i) = 225np(1-p) \end{aligned}$$