CSE 312
Final Exam

Instructions:

• You have 1 hour and 50 minutes to complete the exam.
• Feel free to ask for clarification if something is unclear.
• Please do not turn the page until you are instructed to do so.
• You can use 1 pages of notes during the exam.
• You do not need to simplify any calculations with numbers or simplify sums or integrals.
• If you hit a problem you are finding difficult, I recommend moving on to other problems and then coming back.
• The last 2 pages of the exam are extra credit problems. (There are three extra credit problems in total.) You do not need to do any of them to get a perfect score on the test or a 4.0 in the class (the latter assumes of course that you do well on all your other work). You shouldn’t work on extra credit problems until you are satisfied with what you’ve done on the regular problems. Each extra credit problem is worth 10 points, but as I explained in class, these points mean something different.
• Good luck!
1. (a) **True or False:** The probability of getting 20 heads in 100 independent tosses of a coin that has probability $5/6$ of coming up heads is $(5/6)^{20}(1/6)^{80}$.

(b) **True or False:** Suppose we roll a six-sided fair die twice independently. Then the event that the first roll is 3 and the sum of the two rolls is 6 are independent.

(c) **True or False:** If $X$ and $Y$ are independent random variables, then so are $X^2$ and $Y^2$.

(d) **True or False:** The central limit theorem requires the random variables to be independent.

(e) **True or False:** If $X \sim \text{Unif}(0, 1)$, then $Y = -\ln X$ is an exponential with parameter $\lambda = 1$.

(f) **True or False:** If $X \sim \text{Unif}(-1, 1)$, and we define $Y = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, then $Y \sim \mathcal{N}(0, 1)$.

(g) **True or False:** The expectation-maximization algorithm theoretically converges to the maximum likelihood estimate for the parameters.

(h) **True or False:** Let $A$, $B$ and $C$ be any three events defined with respect to a probability space. Then $\Pr(A \cap B \cap C) = \Pr(A \cap B|C)\Pr(B|C)\Pr(C)$.

(i) **True or False:** Let $A$ be the event that a random 5-card poker hand is a 4 of a kind (i.e. contains 4 cards of 1 rank and 1 card of a different rank) and let $B$ be the event that it contains at least one pair. The events $A$ and $B$ are not independent.

(j) **True or False:** If you flip a fair coin 1000 times, then the probability that there are 800 heads in total is the same as the probability that there are 80 heads in the first 100 flips.

(k) **True or False:** If $N$ is a nonnegative integer valued random variable, then

$$\mathbb{E}
\left[
\binom{N}{2}
\right]
= \left(\mathbb{E}[N]\right)^2.$$
2. Short answer: (No explanations please!)

(a) (4 points) Consider a set $S$ containing $k$ distinct integers. What is the smallest $k$ for which $S$ is guaranteed to have 3 numbers that are the same mod 5?

(b) (3 points) Let $X$ be a random variable that takes values between $-10$ and 10. What is the smallest possible value the variance of $X$ can take?

(c) (5 points) How many ways are there to rearrange the letters in the word KNICKKNACK?

\[
\frac{10!}{4!2!2!}
\]

(d) (5 points) What is the coefficient of $x^6$ in the expansion of $(3x^2 + y)^5$?

\[
\binom{5}{3}3^3y^2
\]

(e) (5 points) Describe the probability mass function of a discrete distribution with mean 10 and variance 9 that takes only 2 distinct values.

\[
X = \begin{cases} 
13 & \text{with prob } \frac{3}{5} \\
7 & \text{with prob } \frac{2}{5} 
\end{cases}
\]

(f) (6 points) Suppose I give you a list of 20 possible questions and tell you that a random subset of 8 of them will be on the upcoming test. If you memorize answers to 15 of them and have no clue how to answer the remaining 5, what is the probability that you will get 6 of the 8 questions on the test right?

\[
\frac{\binom{15}{6}\binom{5}{2}}{\binom{20}{8}}
\]

(g) (8 points) Consider a six-sided die where $Pr(1) = Pr(2) = Pr(3) = Pr(4) = 1/8$ and $Pr(5) = Pr(6) = 1/4$. Let $X$ be the random variable which is the square root of the value showing. (For example, if the die shows a 1, $X$ is 1, if the die shows a 2, $X$ is $\sqrt{2}$, if the die shows a 3, $X = \sqrt{3}$ and so on.) What is the expected value of $X$? (You can leave your answer in the form of a sum and do not need to simplify it.)

\[
\frac{1}{8}(\sqrt{1} + \sqrt{2} + \sqrt{3} + \sqrt{4}) + \frac{1}{4}(\sqrt{5} + \sqrt{6})
\]
(h) (10 points) What is the conditional probability that a random 5-card poker hand is a 4 of a kind (i.e. contains 4 cards of 1 rank and 1 card of a different rank) given that it contains at least one pair?

\[
\frac{\Pr(\text{4 of a kind \& at least one pair})}{\Pr(\text{at least one pair})} = \frac{13 \cdot 12 \cdot 4}{1 - \frac{13}{52} \cdot 4^5}
\]

(i) (10 points) Let \( X \) be a random variable with probability density function

\[
f_X(x) = \begin{cases} 
\frac{x}{4} & \text{if } 1 \leq x \leq 3 \\
0 & \text{otherwise,}
\end{cases}
\]

and let \( A \) be the event \( \{X \geq 2\} \). What is \( E(X|A) \)? (You do not need to evaluate integrals or sums or simplify your answer at all.)

\[
\int_{x \mid A} x f(x) \, dx = \int_{A} x f(x) \, dx = \frac{3}{2} \int_{A} x^2 \, dx
\]

(j) Extra Credit: Give an example of a dependent sequence of random variables \( X_1, X_2, \ldots, X_n \) that all have the same expectation, for which for which the weak law of large numbers does NOT hold.

\[
X_1 = \begin{cases} 
1 & \text{with prob } \frac{1}{2} \\
-1 & \text{with prob } \frac{1}{2}
\end{cases}
\]

\[X_2 = \ldots = X_n = X_1\]

\[E(X_i) = 0 \quad \forall i\]

\[
\frac{X_1 + X_2 + \ldots + X_n}{n} \text{ either } 1 \text{ or } -1
\]
3. (20 points) Consider a boolean formula on \( n \) variables in 3-CNF, that is, conjunctive normal form with 3 literals per clause. This means that it is an “and” of “ors”, where each “or” has 3 literals. Each parenthesized expression (i.e., each “or” of three literals) is called a clause. Here is an example of a boolean formula in 3-CNF, with \( n = 6 \) variables and \( m = 4 \) clauses.

\[
(x_1 \lor x_3 \lor x_5) \land (\neg x_1 \lor \neg x_2 \lor x_6) \land (x_5 \lor \neg x_3 \lor x_4) \land (\neg x_1 \lor x_4 \lor x_5).
\]

- (5 points) What is the probability that \((\neg x_1 \lor \neg x_2 \lor x_3)\) evaluates to true if variable \( x_i \) is set to true with probability \( p_i \), independently for all \( i \)?

\[
1 - p_1 \cdot p_2 \cdot (1 - p_3)
\]

- (10 points) Consider a boolean formula in 3-CNF with \( n \) variables and \( m \) clauses. What is the expected number of satisfied clauses if each variable is set to true independently with probability \( 1/2 \)? A clause is satisfied if it evaluates to true. (In the displayed example above, if \( x_1, \ldots, x_5 \) are set to true and \( x_6 \) is set to false, then all clauses but the second are satisfied.)

\[
\Pr \text{ a clause is satisfied } = \frac{7}{8}
\]

\[
X = \# \text{ satisfied clauses}
\]

\[
E(X) = \frac{7}{8} m \quad \text{by linearity of expectation}
\]

- (5 points) Use Markov’s Inequality to get an upper bound on the probability that the number of satisfied clauses is \( m \).

\[
\Pr(X = m) = \Pr(X \geq m) \leq \frac{7}{8}
\]
4. (20 points) More short answer:

- A bus route has interarrival times that are exponentially distributed with parameter $\lambda = 0.05 \text{min}^{-1}$. What is the probability of waiting an hour or more for a bus?

  \[
  X \text{ waiting time to next bus } \Rightarrow E(X) = \frac{1}{\lambda} = 20 \text{ mins}
  \]

  so $\Pr(X \geq 1 \text{ hr}) = \Pr(X \geq 3 E(X)) \leq \frac{1}{3}

- Suppose the random variable $X$ has a normal distribution with mean 9 and variance 49. What is the probability that $X$ takes on a value of at least 23?

  \[
  1 - \Phi \left( \frac{23 - 9}{\sqrt{49}} \right) = \Phi(-2)
  \]

- The time required for a disk access is modeled using a uniform distribution over the interval 3.5 to 5.5 milliseconds. What is the probability that a randomly selected disk access takes more than 4 milliseconds?

  \[
  \frac{1.5}{2} = \frac{3}{4}
  \]
5. Requests arriving at a certain internet server require an average of $\mu = 100$ milliseconds of CPU time to process.

- Give an upper bound on the probability that an arriving request needs more than 1 second (1000 ms) of CPU time to process.

$$\Pr(X > 10 \mu) \leq \frac{1}{10} \quad \text{(Markov)}$$

- Suppose it is known that the CPU demand for individual requests has a standard deviation of $\sigma = 100$ milliseconds. Give an improved upper bound on the probability that an arriving request needs more than 1 second (1000 ms) of CPU time to process.

$$\Pr(|X - \mu| \geq 9\sigma) \leq \frac{1}{81} \quad \text{(Chebyshev)}$$

- Suppose it is known that the CPU demand for individual requests consists of 100 separate, independent steps, each of which averages 1 millisecond with a standard deviation of 1 millisecond. What is the mean CPU demand (in total) for a request? Its standard deviation? What is the approximate form of the distribution?

\[ X \approx \text{normal} \quad \mu = 100 \cdot 1\text{ms} = 100\text{ms} \]
\[ \sigma^2 = 100 \cdot (1\text{ms})^2 = 100\text{ms}^2 \quad \sigma = 10\text{ms} \]

- Continuing the previous part of the question, estimate the probability that a request requires more than 1 second of CPU time. Is this a smaller probability than the bound given in the second part of this problem?

\[ z = \frac{(1000 - \mu)}{\sigma} = 90 \quad \text{so approx prob of requiring more than 1 sec of CPU time is} \quad 1 - \Phi(90) = 0 \]

(significant underestimate)
6. You draw $n$ independent samples $X_1, X_2, \ldots, X_n$ from a Poisson distribution with unknown parameter $\lambda$. What is the likelihood function corresponding to these observations?

See daily problem, March 4

• What is the maximum likelihood estimator, $\hat{\lambda}$, for the unknown parameter $\lambda$?
7. Extra Credit: Let $X_1, X_2, \ldots, X_N$ be independent, identically distributed random variables, where $N$ itself (that is, the number of random variables) is a random variable. Assuming $X_i$ and $N$ are also independent for each $i$, prove that

$$E\left( \sum_{1 \leq i \leq N} X_i \right) = E(N)E(X_1).$$

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

Linearity of expectation and independence

\[
\begin{align*}
&= \sum_{n=0}^{\infty} n E(X_1) Pr(N = n) \\
&= E(X_1) \sum_{n=0}^{\infty} n Pr(N = n) \\
&= E(X_1) E(N)
\end{align*}
\]