Section 10: Final Review

1. True or False?

(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 discrete, 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:** Let A, B and C be any three events defined with respect to a probability space. Then $\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A \cap B|C)\mathbb{P}(B|C)\mathbb{P}(C)$.

(f) **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.

(g) **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.

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

$$\mathbb{E}\left[\binom{N}{2}\right] = \binom{\mathbb{E}[N]}{2}.$$

2. Short answer

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

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

(d) I toss n balls into n bins uniformly at random. What is the expected number of bins with exactly k balls in them?

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

(f) 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? (Leave your answer in the form of a numerical sum; do not bother simplifying it.)

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

(h) How many different ways are there to select 3 dozen indistinguishable colored roses if red, yellow, pink, white, purple and orange roses are available?

(i) Two identical 52-card decks are mixed together. How many distinct permutations of the 104 cards are there?

3. Random boolean formulas

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.

$$(x_1 \vee x_3 \vee x_5) \wedge (\neg x_1 \vee \neg x_2 \vee x_6) \wedge (x_5 \vee \neg x_3 \vee x_4) \wedge (\neg x_1 \vee x_4 \vee x_5).$$

(a) What is the probability that $(\neg x_1 \vee \neg x_2 \vee x_3)$ evaluates to `true` if variable x_i is set to `true` with probability p_i, independently for all i?

(b) 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.)

4. Biased coin flips

We flip a biased coin with probability p of getting heads until we either get heads or we flip the coin three times. Thus, the possible outcomes of this random experiment are $< H >, < T, H >, < T, T, H >$ and $< T, T, T >$.

(a) What is the probability mass function of X, where X is the number of heads. (Notice that X is 1 for the first three outcomes, and 0 in the last outcome.)

(b) What is the probability that the coin is flipped more than once?

(c) Are the events "there is a second flip and it is heads" and "there is a third flip and it is heads" independent? Justify your answer.

(d) Given that we flipped more than once and ended up with heads, what is the probability that we got heads on the second flip? (No need to simplify your answer.)

5. Bitcoin users

There is a population of n people. The number of Bitcoin users among these n people is i with probability p_i, where, of course, $\sum_{0 \leq i \leq n} p_i = 1$. We take a random sample of k people from the population (without replacement). Use Bayes Theorem to derive an expression for the probability that there are i Bitcoin users in the population conditioned on the fact that there are j Bitcoin users in the sample. Let B_i be the event that there are i Bitcoin users in the population and let S_j be the event that there are j Bitcoin users in the sample. Your answer should be written in terms of the p_ℓ's, i, j, n and k.

6. Investments

You are considering three investments. Investment A yields a return which is X dollars where X is Poisson with parameter 2. Investment B yields a return of Y dollars where Y is Geometric with parameter 1/2. Investment C yields a return of Z dollars which is Binomial with parameters $n = 20$ and $p = 0.1$. The returns of the three investments are independent.

(a) Suppose you invest simultaneously in all three of these possible investments. What is the expected value and the variance of your total return?

(b) Suppose instead that you choose uniformly at random from among the 3 investments (i.e., you choose each one with probability 1/3). Use the law of total probability to write an expression for the probability that the return is 10 dollars. Your final expression should contain numbers only. No need to simplify your answer.

7. Another continuous r.v.

The density function of X is given by

$$f(x) = \begin{cases} a + bx^2 & \text{when } 0 \leq x \leq 1 \\ 0 & \text{otherwise.} \end{cases}$$

If $E(X) = \frac{3}{5}$, find a and b.

8. Extended Family Portrait

A group of n families, each with m members, are to be lined up for a photograph. In how many ways can the nm people be arranged if members of a family must stay together?

9. Poisson CLT practice

Suppose $X_1, \ldots, X_n$ are iid $\mathsf{Poisson}(\lambda)$ random variables, and let $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, the sample mean. How large should we choose n to be such that $\mathbb{P}(\frac{\lambda}{2} \leq \overline{X}_n \leq \frac{3\lambda}{2}) \geq 0.99$? Use the CLT and give an answer involving $\Phi^{-1}(\cdot)$. Then evaluate it exactly when $\lambda = 1/10$ using the Φ table on the last page.

10. Law of Total Probability Review

(a) (Discrete version) Suppose we flip a coin with probability U of heads, where U is equally likely to be one of $\Omega_U = \{0, \frac{1}{n}, \frac{2}{n}, ..., 1\}$ (notice this set has size $n+1$). Let H be the event that the coin comes up heads. What is $\mathbb{P}(H)$?

(b) (Continuous version) Now suppose $U \sim \mathsf{Uniform(0,1)}$ has the *continuous* uniform distribution over the interval $[0, 1]$. What is $\mathbb{P}(H)$?

(c) Let's generalize the previous result we just used. Suppose E is an event, and X is a continuous random variable with density function $f_X(x)$. Write an expression for $\mathbb{P}(E)$, conditioning on X.

11. A Red Poisson (From section 9 handout)

Suppose that $x_1, \ldots, x_n$ are i.i.d. samples from a Poisson(θ) random variable, where θ is unknown. Find the MLE of θ.

12. Tail bounds

Suppose $X \sim \mathsf{Geo}(1/4)$. We will bound $\mathbb{P}(X \geq 6)$ using the tail bounds we've learned, and compare this to the true result.

(a) Give an upper bound for this probability using Markov's inequality. Why can we use Markov's inequality?

(b) Give an upper bound for this probability using Chebyshev's inequality. You may have to rearrange algebraically and it may result in a weaker bound.

(c) Give the exact probability.