Q1 Probability True/False
2 Points

Q1.1
0.5 Points

Let $(\Omega, \mathbb{P})$ be any probability space. For all $\omega \in \Omega$, $0 < \mathbb{P}[\omega] < 1$.

- ○ True
- ◉ False

Explanation

Non-negativity: for all $\omega \in \Omega$, $0 \leq \mathbb{P}[\omega] \leq 1$.

Q1.2
0.5 Points

For any two events $A, B \subseteq \Omega$ in the same probability space, $\mathbb{P}[A \cup B] = \mathbb{P}[A] + \mathbb{P}[B] - \mathbb{P}[A]\mathbb{P}[B]$.

- ○ True
- ◉ False

Explanation

The equation above is true when A and B are independent events. But A and B are any two events, not necessarily two independent events.

Q1.3
0.5 Points

For any three mutually exclusive events $A_1, A_2, A_3 \subseteq \Omega$ in the same probability space, $\mathbb{P}[A_1 \cup A_2 \cup A_3] = \mathbb{P}[A_1] + \mathbb{P}[A_2] + \mathbb{P}[A_3] - \mathbb{P}[A_1 \cap A_2] - \mathbb{P}[A_1 \cap A_3] - \mathbb{P}[A_2 \cap A_3]$.

- (●) True
- ( ) False

Explanation

If A_1, A_2, A_3 are mutually exclusive events, then $\mathbb{P}[A_1 \cup A_2 \cup A_3] = \mathbb{P}[A_1] + \mathbb{P}[A_2] + \mathbb{P}[A_3]$. The intersection probabilities are all 0 because of mutual exclusivity, so the equation still holds.

Q1.4
0.5 Points

Consider a Markov chain over a finite state space S with transition matrix P.

If the chain has an invariant distribution, then it is irreducible.

- ( ) True
- (●) False

Explanation

Consider the case where $S = \{1, 2\}$ and $P = \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}$. This is not irreducible because you cannot go from state 2 to state 1. But $\pi = \begin{bmatrix} 0 & 1 \end{bmatrix}$ is an invariant distribution for the Markov chain.

Q2 Identically Distributed

2 Points

Let X, Y be random variables with the same distribution.

Q2.1

0.5 Points

For all k, $\mathbb{P}[X \leq k] = \mathbb{P}[Y \leq k]$.

- (●) True
- ( ) False

Explanation

Two random variables have the same distribution iff they have the same CDF. There are also other similar equations that are true because X and Y have the same distribution, such as:

- For all k, $\mathbb{P}[X = k] = \mathbb{P}[Y = k]$
- For all k, $\mathbb{P}[X < k] = \mathbb{P}[Y < k]$
- For all k, $\mathbb{P}[X \geq k] = \mathbb{P}[Y \geq k]$

Q2.2

0.5 Points

$\mathbb{E}[X + Y] = \mathbb{E}[2X]$.

- (●) True
- ( ) False

Explanation

By linearity of expectation, $\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$. Since X and Y have the same distribution, they have the same expectation, so $\mathbb{E}[X] + \mathbb{E}[Y] = \mathbb{E}[X] + \mathbb{E}[X] = 2\mathbb{E}[X] = \mathbb{E}[2X]$.

Q2.3
0.5 Points

$\mathrm{Var}(X+Y) = 2\mathrm{Var}(X)$.

- ○ True
- ◉ False

Explanation

$\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{cov}(X,Y)$. Since X and Y have the same distribution, they have the same variance, so $\mathrm{Var}(X) + \mathrm{Var}(Y) - 2\mathrm{cov}(X,Y) = \mathrm{Var}(X) + \mathrm{Var}(X) - 2\mathrm{cov}(X,Y) = 2\mathrm{Var}(X) - 2\mathrm{cov}(X,Y)$.

But $\mathrm{cov}(X,Y)$ is not necessarily 0. (X and Y being independent would imply that their covariance is zero, but the question did not assume that X and Y were independent.)

Q2.4
0.5 Points

$\operatorname{cov}(X, Y) = \operatorname{Var}(X)$.

○ True

◉ False

Explanation

This equation is true when $X = Y$ (they are the same random variable). But just because X and Y have the same distribution doesn't mean that they are the same random variable: for a given outcome, X and Y can assign different values.

A concrete example that distinguishes "identical distribution" from "the same random variable": suppose you 2 fair six-sided dice. Let X be the outcome of the first die, and Y be the outcome of the second die.

X and Y have the same distribution. For $k = 1, 2, 3, 4, 5, 6$: $\mathbb{P}[X = k] = \mathbb{P}[Y = k] = \frac{1}{6}$. But they aren't the same random variable: for instance, for the outcome $(1, 2)$, $X = 1$ but $Y = 2$.

Accordingly, $\operatorname{cov}(X, Y) = \operatorname{Var}(X)$ is false. X and Y are independent, so their covariance is zero, but $\operatorname{Var}(X) = \frac{35}{12}$.

Q3 Famous Distributions

5 Points

Q3.1

1 Point

Let $X_1, X_2, \ldots, X_5$ be i.i.d. $\text{Bernoulli}(p)$ random variables, and $X = \sum_{i=1}^{5} X_i$.

Let $Y = \sum_{i=1}^{3} X_i$.

What is $\mathbb{P}[X = 3 \mid Y = 2]$?

- ○ $\binom{5}{3} p^3 (1-p)^2$
- ○ p^3
- ○ $3p(1-p)^2$
- ◉ $2p(1-p)$

Explanation

Given that $Y = 2$, the only way to get $X = 3$ is for exactly one of X_4, X_5 to be 1 and for the other to be 0. There are 2 ways to choose which of X_4, X_5 is 1, p chance that it is indeed 1, and $(1-p)$ chance for the other RV to be 0.

Q3.2
1 Point

Let $X_1, X_2, \ldots, X_n$ be i.i.d. $\text{Bernoulli}(p)$ random variables for some fixed p, and let $S_n = \sum_{i=1}^n X_i$.

As $n \to \infty$, which of the following statements is correct?

- ○ $S_n \to \text{Poisson}(np)$
- ○ $S_n \to \text{Binomial}(n, p)$
- ○ $\frac{S_n - np}{\sqrt{np(1-p)}} \to \text{Poisson}(p)$
- ◉ $\frac{S_n - np}{\sqrt{np(1-p)}} \to N(0, 1)$

Explanation

This is the result of the Central Limit Theorem. As n approaches infinity, the distribution of the standardized sum of i.i.d. random variables converges to a standard normal distribution.

Note that for the Poisson distribution as a limit of the Binomial distribution, the probability of success $p \to 0$ as $n \to \infty$. But in this question, p is fixed.

Q3.3
1 Point

Let $X_1, X_2, \ldots X_n$ be i.i.d. $\text{Exp}(\lambda)$. If $X = \min(X_1, X_2, \ldots, X_n)$, what is $\mathbb{P}[X \geq n]$?

- ○ $\frac{1}{n^2\lambda}$
- ○ $\frac{1}{n^2\lambda^2(n-n\lambda)^2}$
- ○ $e^{-n\lambda}$
- ◉ $e^{-n^2\lambda}$

Explanation

Notice that $\mathbb{P}[X \geq n] = \mathbb{P}[(X_1 \geq n) \cap (X_2 \geq n) \cap \ldots (X_n \geq n)]$, since all the X_i must be at least n for their minimum to be at least n. Because the X_i are i.i.d, this probability becomes $(\mathbb{P}[X_1 \geq n])^n = (e^{-\lambda n})^n = e^{-n^2\lambda}$.

Q3.4
1 Point

Let $X \sim \text{Geometric}(p)$ distribution. Given $X = x$, let U be a uniform random variable on $[0, x]$.

What is $\mathbb{E}[U]$?

- ○ $\frac{x}{2}$
- ○ $\frac{1}{p}$
- ◉ $\frac{1}{2p}$
- ○ The European Union

Explanation

We'll use total expectation. $\mathbb{E}[U \mid X = x] = \frac{x}{2}$, so $\mathbb{E}[U \mid X] = \frac{X}{2}$. Thus, $\mathbb{E}[U] = \mathbb{E}[\mathbb{E}[U \mid X]] = \mathbb{E}[\frac{X}{2}] = \frac{1}{2}\mathbb{E}[X]$. Since $X \sim \text{Geometric}(p)$, $\mathbb{E}[X] = \frac{1}{p}$, so $\mathbb{E}[U] = \frac{1}{2p}$.

Q3.5
1 Point

Let Z_1 be a standard normal random variable. Given $Z_1 = z_1$, let $Z_2 \sim N(z_1, 2)$. What is the distribution of $\mathbb{E}[Z_2 \mid Z_1]$?

- ◉ $N(0, 1)$
- ○ 0
- ○ $N(0, 2)$
- ○ $N(z_1, 2)$

Explanation

$\mathbb{E}[Z_2 \mid Z_1 = z_1] = z_1$. Hence, $\mathbb{E}[Z_2 \mid Z_1] = Z_1$, which is also standard normal.

Q4 Some More Applications
3 Points

IP addresses have two formats: IPv4 and IPv6.

- IPv4 addresses are of the form `x.x.x.x`, where `x` can be any integer from 0 to 255.
- IPv6 addresses are of the form `xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx:xxxx`, where `x` can be any hexadecimal digit from 0 to f (there are 16 possible values for `x`).

Suppose a hacker guesses an IP address uniformly at random, once per nanosecond.

Q4.1
1 Point

What is the expected amount of time until the hacker guesses your IP address if you're using (a) IPv4; (b) IPv6? Assume your IP address stays the same.

- ○ IPv4: 2^{1024} nanoseconds; IPv6: 2^{512} nanoseconds
- ○ IPv4: 2^{1024} nanoseconds; IPv6: 2^{128} nanoseconds
- ○ IPv4: 2^{32} nanoseconds; IPv6: 2^{512} nanoseconds
- ◉ IPv4: 2^{32} nanoseconds; IPv6: 2^{128} nanoseconds

Explanation

There are $256^4 = 2^{32}$ different IPv4 addresses, so the probability that the hacker guesses it right is $\frac{1}{2^{32}}$. Let X is a random variable denoting the number of nanoseconds until the hacker guesses it right, so $X \sim \mathrm{Geometric}(\frac{1}{2^{32}})$. We want to find $\mathbb{E}[X]$, which is $\frac{1}{\frac{1}{2^{32}}} = 2^{32}$. For context, 2^{32} nanoseconds is about 5 seconds.

Similarly, there are $16^{32} = 2^{128}$ different IPv6 addresses. Using the same reasoning as above, the hacker expects to guess your IPv6 address in 2^{128} milliseconds, which is about 10^{22} years. (Get good, hacker.)

Q4.2
1 Point

Approximately how long will it take before the hacker guesses a repeat IP address with probability $\frac{1}{2}$?

- (selected) IPv4: 2^{16} nanoseconds; IPv6: 2^{64} nanoseconds
- IPv4: 2^{31} nanoseconds; IPv6: 2^{127} nanoseconds
- IPv4: 2^{512} nanoseconds; IPv6: 2^{256} nanoseconds
- IPv4: 2^{1023} nanoseconds; IPv6: 2^{511} nanoseconds

Explanation

This is the birthday paradox problem, where the number of bins n is the number of IP addresses (IPv4: 2^{32}; IPv6: 2^{128}). The number of guesses made before a guess is repeated with probability $\frac{1}{2}$ is approximately $\sqrt{n}$. So for IPv4, $\sqrt{2^{32}} = 2^{16}$, and for IPv6, $\sqrt{2^{128}} = 2^{64}$.

Q4.3
1 Point

We can also view IP addresses as binary strings.

A different hacker comes up with a different way to guess IP addresses: rather than guessing an entire IP address uniformly at random each nanosecond, they pick 1 bit uniformly at random from the most recently guessed address, and flip it to form the new guess.

Suppose this hacker's first IP address ends in `...101`. Let E be the event that they guess an IP address whose last 3 bits contain exactly two consecutive 0s (e.g., `...100`) before they guess an IP address whose last 3 bits contain exactly two consecutive 1s (e.g., `...110`).

Select the true statement about $\mathbb{P}[E]$.

- ○ $\mathbb{P}[E] < \frac{1}{2}$
- ○ $\mathbb{P}[E] = \frac{1}{2}$
- ● $\mathbb{P}[E] > \frac{1}{2}$

Explanation

We can model this as a Markov chain problem, whose states are the vertices on a 3-dimensional hypercube (representing the last 3 bits of the current guess). The transition probabilities are $\frac{1}{3}$ between adjacent vertices and 0 otherwise. We only make a transition along the hypercube when the last 3 bits change.

We want to find the probability that we're in a state in $A = \{001, 100\}$ before we're in a state in $B = \{110, 011\}$, given that we're in state 101.

$\alpha(101) = \frac{1}{3}\alpha(001) + \frac{1}{3}\alpha(111) + \frac{1}{3}\alpha(100) = \frac{2}{3} + \frac{1}{3}\alpha(101) > \frac{1}{2}$.

If you're curious about the exact value of $\alpha(101)$: $\alpha(111) = \frac{1}{3}\alpha(011) + \frac{1}{3}\alpha(101) + \frac{1}{3}\alpha(110) = \frac{1}{3}\alpha(101)$. Plugging this into $\alpha(101)$ above and solving gives us $\alpha(101) = \frac{3}{4}$.