Probability and distributions in WBS Primer - Quiz notes

What is not an example of a random process? User clicks on a website in August 2022.

An example of a non-random process would be "User clicks on a website in August 2022" when *the clicking behavior is influenced by specific factors rather than being completely random*. Let's break this down:

2. **External Influences**: The user's clicking behavior could also be influenced by external events. For example, if there's a major sports event in August 2022, a user interested in sports is more likely to click on related websites. This decision is influenced by the event, not by chance.

In summary, *a non-random process involves actions or events that have identifiable causes or patterns, as opposed to random processes where outcomes are unpredictable and not influenced by prior events or patterns*.

Having rolled a die twice, what are possible events? All of them

All of them

What is a useful property of the normal distribution? Its shape can be described using only μ and σ.

A useful property of the normal distribution is indeed its shape, which is fully determined by just two parameters: the mean (μ) and the standard deviation (σ). This characteristic makes the normal distribution uniquely manageable and informative in statistics and probability theory. Here's how these two parameters define the distribution:

1. **Mean (μ)**: The mean is the central point of the normal distribution. In a graph of the distribution, it's the location of the peak. It represents the average or expected value of the data. *In a normal distribution, about 50% of the values lie below the mean and 50% above it. So the mean, the median and the mode are equal*
2. **Standard Deviation (σ)**:
   > The power of a normal distribution derives from the fact that we know roughly what proportion of observations will lie within one standard deviation above or below the mean (68 percent); what proportion of observations will lie within two

> standard deviations above or below the mean (95 percent); and so on. — location: [2372]

The ability to describe the shape of the normal distribution with only the mean and standard deviation makes it a fundamental tool in statistical analysis and probability theory.

What do empirical distributions tend to look like? Skewed

Empirical distributions *represent the actual data we observe or collect from experiments or real-world observations*. When we talk about these distributions being "skewed," we're describing how they're not perfectly balanced or symmetrical.

Which of these statements about probability density functions is true? It covers all possible events on the x-axis.

A **P**robability **D**ensity **F**unction, used in the context of continuous random variables, does indeed cover all possible events on the x-axis. However, there are a few key points to understand:

1. **Range of Possible Events:** *The x-axis in a PDF represents the range of all possible outcomes for a continuous random variable. For instance, if the variable is the height of adults, the x-axis would cover all possible heights.*
2. **Probability Density, Not Probability:** Unlike probability mass functions for discrete random variables, a PDF does not directly give you the probability of an event. Instead, it provides a density. The probability of an event occurring within a specific range is found by calculating the area under the curve of the PDF over that range.
3. **Total Area Under the Curve:** For a PDF, the total area under the curve across all possible values (the entire x-axis) is always equal to 1. This signifies that the probability of some event happening (within the range of possible events) is certain.
4. **Individual Points:** In a PDF, the probability of the random variable taking on any single, specific value (like exactly 170.5 cm in height) is technically zero because the range of values is continuous. We only deal with ranges or intervals for probability in continuous distributions.

Which of the following statements about random variables is not true? They are usually denoted by lowercase letters

The statement that "random variables are usually denoted by lowercase letters" is not true. *In the context of statistics and probability theory, random variables are* *conventionally denoted by uppercase letters*. For example, we might use $ (X, Y,)$ or (Z) to represent random variables.

Random variables are essentially variables that take on different values based on the outcomes of a random process or experiment. There are two main types:

1. **Discrete Random Variables:** These variables take on a countable number of distinct values. For instance, the number of heads in ten coin flips is a discrete random variable.
2. **Continuous Random Variables:** These variables can take on any value within a given range. The exact height of randomly selected individuals would be an example of a continuous random variable.

In contrast, specific outcomes or values that the random variables can take are often denoted by lowercase letters. For example, we might say (X = x) to indicate that the random variable (X) takes the specific value (x).

What is true about probability? The probabilities of all possible outcomes add up to 1.

Yes, that's correct. *One of the fundamental principles of probability is that the sum of the probabilities of all possible outcomes of a random experiment adds up to 1.* This principle is based on the concept of a probability space, which encompasses all possible outcomes of a given random process.

Here are a few key points to understand this concept:

1. **Definition of Probability:** Probability is a measure that quantifies the likelihood of an event occurring. It ranges from 0 (impossible event) to 1 (certain event).
2. **All Possible Outcomes:** When considering all possible outcomes of a random process, it's understood that one of these outcomes must occur. For example, when flipping a fair coin, there are two possible outcomes: heads or tails. The probability of getting heads plus the probability of getting tails equals 1.
3. **Discrete and Continuous Cases:** This principle applies both in discrete cases (like flipping a coin or rolling a dice, where there are a finite number of outcomes) and in continuous cases (like measuring the height of a person, where the outcome can be any value within a range). In the continuous case, the integral of the probability density function over the entire range of possible values equals 1.
4. **Mutually Exclusive and Exhaustive Events:** In a scenario where events are mutually exclusive (only one event can occur at a time) and exhaustive (covering all possible

outcomes), the sum of their probabilities is 1. This is often visualized in probability trees or Venn diagrams in more complex scenarios.

This principle ensures that probability is a consistent and reliable measure in various fields, from basic experiments to complex statistical analyses.

Which of the following are disjoint events? Toss a coin. E = {Heads, Tail}

Disjoint events, also known as mutually exclusive events, are events that cannot happen at the same time. In the context of your example, where you toss a coin once, the events in set $(E = \{\text{Heads}, \text{Tails}\})$ are indeed disjoint events.

Here's why:

- If you toss a coin, it can only land on either heads or tails, but not both at the same time.
- The occurrence of one event (e.g., the coin landing on heads) completely rules out the possibility of the other event (the coin landing on tails) occurring in that single toss.

In probability theory, when two events are disjoint, the probability of both occurring simultaneously is zero. For a coin toss, the probability that you get both heads and tails in one toss is 0, which confirms that they are disjoint events.

Which of the following observations can be explained by the law of large numbers? Coin tosses: probability of Heads = 0.3. after 10, and = 0.41 after 100.

In your example, the expected theoretical probability of getting heads is 0.3. After just 10 tosses, you might not see a result very close to 0.3 due to the small sample size; randomness has a more significant impact in small samples, leading to greater variability. However, as you increase the number of tosses to 100, the observed probability of getting heads becomes closer to the expected probability of 0.3. *This illustrates the Law of Large Numbers, where the average outcome over many trials tends to stabilize and converge towards the expected value.*

It's important to note that the Law of Large Numbers doesn't guarantee that the observed probability will exactly match the expected probability, especially not in every individual experiment. Rather, it indicates a trend over time and many repetitions towards the expected value. Also, the observed probability of 0.41 after 100 tosses, while closer *to 0.3 than the result after 10 tosses, still shows some deviation, which could be due to random variation or other factors influencing the outcomes.*

Extra notes:

> But probability triumphs in the end. An important theorem known as the law of large numbers tells us that as the number of independent trials increases, the average of the outcomes will get closer and closer to its expected value. — location: [1436]