

Understanding Recidivism Rate

You said:



 Uploaded image

I have a problem with the solution, related to Data Analysis, as shown in the picture.
Please explain the context of the problem.
Please extract all specific terms, math expressions, and notations.
Then explain each with an analogy and a quantitative example.

ChatGPT said:



Context of the Problem:

The problem involves determining the expected recidivism rate (the likelihood that a released prisoner will commit additional crimes) based on various factors such as the length of the original sentence, the distance of the prison from the hometown, and the gender of the offender. The recidivism rate is modeled using a mathematical formula that incorporates these variables.

Specific Terms, Math Expressions, and Notations:

1. Recidivism Rate ($P(R)$):

- The probability that a released prisoner will commit additional crimes.

2. Length of Original Sentence (S):

- The length of time the prisoner was initially sentenced to serve, assumed to be uniformly distributed between 1 and 15 years ($U(1, 15)$).

3. Distance from Hometown (D):

- The distance between the prison and the prisoner's hometown, assumed to follow an exponential distribution with a rate parameter of 0.05 ($\exp(0.05)$).

4. Sex of Offender (M):

- A binary variable where $M = 1$ if the offender is male and $M = 0$ if female. It is given that 80% of offenders are male.

5. Formula for Recidivism Rate:

$$P(R) = \frac{S}{32} + \frac{D}{40} + 0.1M$$

6. Expected Value (\mathbb{E}):

- Represents the average or mean value of a random variable over its possible outcomes. The expected value of S , D , and M are calculated to find the overall expected recidivism rate.

Explanation with Analogies and Quantitative Examples:

1. Length of Original Sentence (S):

- **Analogy:** Imagine you are rolling a 15-sided die, where each side represents a possible sentence length from 1 to 15 years. The average outcome (expected value) after many rolls would be the middle value, which is 8.
- **Example:** If S is uniformly distributed between 1 and 15, its expected value ($\mathbb{E}[S]$) is:

$$\mathbb{E}[S] = \frac{1 + 15}{2} = 8 \text{ years}$$

- **Contribution to Recidivism:** The term $\frac{S}{32}$ represents the influence of the sentence length on the recidivism rate. With $\mathbb{E}[S] = 8$, this contribution is $\frac{8}{32} = 0.25$.

2. Distance from Hometown (D):

- **Analogy:** Suppose you're drawing random distances from a long road, where short distances are more common than long ones. The average distance would be relatively close to the starting point.
- **Example:** If D follows an exponential distribution with a rate parameter of 0.05, its expected value ($\mathbb{E}[D]$) is:

$$\mathbb{E}[D] = \frac{1}{0.05} = 20 \text{ units (e.g., miles)}$$

- **Contribution to Recidivism:** The term $\frac{D}{40}$ represents the influence of distance on recidivism. With $\mathbb{E}[D] = 20$, this contribution is $\frac{20}{40} = 0.5$.

3. Sex of Offender (M):

- **Analogy:** Consider a bag with 10 marbles, where 8 are red (representing males) and 2 are blue (representing females). If you draw one marble, the chance of getting a red one (a male offender) is 80%.
- **Example:** Given that 80% of offenders are male, the expected value of M ($\mathbb{E}[M]$) is:

$$\mathbb{E}[M] = 0.8$$

- **Contribution to Recidivism:** The term $0.1M$ reflects the gender impact on recidivism. With $\mathbb{E}[M] = 0.8$, this contribution is $0.1 \times 0.8 = 0.08$.

Calculating the Expected Recidivism Rate:

The overall expected recidivism rate combines these contributions:

$$\mathbb{E}[R] = \frac{\mathbb{E}[S]}{32} + \frac{\mathbb{E}[D]}{40} + 0.1 \times \mathbb{E}[M]$$

Substituting the expected values:

$$\mathbb{E}[R] = \frac{8}{32} + \frac{20}{40} + 0.1 \times 0.8 = 0.25 + 0.5 + 0.08 = 0.83 \text{ or } 83\%$$

Thus, the expected recidivism rate is 83%, meaning that, on average, 83 out of 100 released prisoners would be expected to re-offend based on the given model.

You said:



Please explain more the following:
uniform distribution $U(1,15)$,

exponential distribution $\exp(0.05)$.

ChatGPT said:



Certainly! Let's break down the concepts of a **uniform distribution** and an **exponential distribution** with specific focus on the examples provided: $U(1, 15)$ and $\exp(0.05)$.

1. Uniform Distribution $U(1, 15)$

Explanation:

- **Uniform Distribution:** A uniform distribution is a type of probability distribution where all outcomes are equally likely. In other words, each value within the specified range has the same probability of occurring.
- $U(1, 15)$: This notation indicates a uniform distribution over the interval from 1 to 15. It means that any value between 1 and 15 is equally likely to be selected.

Analogy:

- **Analogy:** Imagine a perfectly fair 15-sided die where each side is numbered from 1 to 15. When you roll the die, each number (1, 2, ..., 15) has an equal chance of landing face up. There's no preference for any particular number, so the probability of getting any specific number is the same as any other.

Example:

- **Example:** If you were to randomly select a value from $U(1, 15)$, the expected or average value (also called the mean) would be the midpoint of 1 and 15. Mathematically, the expected value $\mathbb{E}[S]$ is calculated as:

$$\mathbb{E}[S] = \frac{1 + 15}{2} = 8$$

So, if you repeated this random selection process many times, on average, you would get a value of 8.

Graphical Representation:

- **Graph:** The graph of a uniform distribution $U(1, 15)$ would be a flat, horizontal line between 1 and 15. The height of the line represents the probability density, and since every number between 1 and 15 is equally likely, the line is constant across the interval.

2. Exponential Distribution $\exp(0.05)$

Explanation:

- **Exponential Distribution:** An exponential distribution is often used to model the time between events in a process that happens continuously and independently at a constant average rate. It's a probability distribution characterized by its rate parameter, denoted here as $\lambda = 0.05$.
- $\exp(0.05)$: This notation indicates an exponential distribution with a rate parameter $\lambda = 0.05$. The rate parameter controls how quickly the probabilities decrease as you move away from zero. A smaller λ results in a slower decay, meaning larger values are more probable than with a larger λ .

Analogy:

- **Analogy:** Think of a situation where you're waiting for a bus, and the buses arrive randomly, but on average, one bus arrives every 20 minutes. The time you wait for the next bus can be modeled by an exponential distribution. If $\lambda = 0.05$, it would imply that the average waiting time (or distance, in your context) is 20 units (since $\frac{1}{\lambda} = 20$).

Example:

- **Example:** In the context of the distance D , if the distance from the prison to the hometown follows $\exp(0.05)$, the expected or average distance can be calculated as:

$$\mathbb{E}[D] = \frac{1}{\lambda} = \frac{1}{0.05} = 20 \text{ units (e.g., miles, kilometers)}$$

This means that, on average, the distance between the prison and the hometown is 20 units.

Graphical Representation:

- **Graph:** The graph of an exponential distribution $\exp(0.05)$ would start at its highest point at 0 and then rapidly decrease as you move to the right. The curve is steep at the beginning and flattens out as it extends further, reflecting that small values are more probable, but larger values can still occur with decreasing likelihood.

Summary:

- **Uniform Distribution** $U(1, 15)$: Models a situation where every outcome between 1 and 15 is equally likely. The average value is 8.
- **Exponential Distribution** $\exp(0.05)$: Models a situation where smaller values are more likely, but larger values can still occur. The average value is 20 units.

These distributions are used in the problem to model the length of the sentence and the distance from the prison to the hometown, which are then used to calculate the expected recidivism rate.

Is this conversation helpful so far?

ChatGPT can make mistakes. Check important info.