Python Examples for Random Variables

Discrete Random Variable

Listing 1: PMF and Expected Value for a Discrete Random Variable

```
# PMF and Expected Value for a Discrete Random Variable
outcomes = [0, 1, 2, 3]
probabilities = [0.1, 0.5, 0.3, 0.1]

# Expected Value
expected_value = sum(x * p for x, p in zip(outcomes, probabilities))
print(f"Expected Value: {expected_value}")

# Variance
variance = sum((x - expected_value)**2 * p for x, p in zip(outcomes,
    probabilities))
print(f"Variance: {variance}")
```

Continuous Random Variable

Listing 2: PDF and CDF for a Continuous Random Variable

```
from scipy.stats import norm

# PDF and CDF for a Continuous Random Variable
mean, std_dev = 0, 1  # Standard normal distribution
x = 1.0

# PDF
pdf_value = norm.pdf(x, mean, std_dev)
print(f"PDF at X={x}: {pdf_value}")

# CDF
cdf_value = norm.cdf(x, mean, std_dev)
print(f"CDF at X={x}: {cdf_value}")
```

Joint Distribution

Listing 3: Joint PMF for Discrete Random Variables

```
import numpy as np

# Joint PMF for Discrete Random Variables
joint_pmf = np.array([[0.1, 0.2], [0.3, 0.4]])  # 2x2 table for P(X, Y)

# Marginal Distribution
marginal_X = joint_pmf.sum(axis=1)  # Sum over Y
marginal_Y = joint_pmf.sum(axis=0)  # Sum over X

print(f"Marginal Distribution for X: {marginal_X}")
print(f"Marginal Distribution for Y: {marginal_Y}")
```

Detailed Explanation of Random Variables

A **random variable** is a numerical description of the outcomes of a random experiment. It is a key concept in probability and statistics and is used to model real-world phenomena where outcomes are uncertain.

Types of Random Variables

1. **Discrete Random Variables**:
   - Take a finite or countable set of values.
   - **Example**: Number of heads in 3 coin tosses ($X = 0, 1, 2, 3$).
2. **Continuous Random Variables**:
   - Take any value in a continuous range.
   - **Example**: Height of people in a population ($X \in [0, \infty)$).

Key Properties of Random Variables

1. Probability Distribution

The **probability distribution** of a random variable describes how probabilities are assigned to its possible values.

- **Discrete Random Variable**:
  - Described by the **Probability Mass Function (PMF)**, $P(X = x)$:
  $$P(X = x) \geq 0 \quad \text{and} \quad \sum_x P(X = x) = 1$$
  - **Example**: Tossing a fair coin twice ($X =$ Number of Heads):
  $$P(X = 0) = \frac{1}{4}, \quad P(X = 1) = \frac{1}{2}, \quad P(X = 2) = \frac{1}{4}$$
- **Continuous Random Variable**:
  - Described by the **Probability Density Function (PDF)**, $f_X(x)$:
  $$f_X(x) \geq 0 \quad \text{and} \quad \int_{-\infty}^{\infty} f_X(x)dx = 1$$
  - The probability of X in a range $[a, b]$ is:
  $$P(a \leq X \leq b) = \int_a^b f_X(x)dx$$
  - **Example**: Standard normal distribution:
  $$f_X(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

2. Cumulative Distribution Function (CDF)

The CDF describes the probability that a random variable X is less than or equal to a specific value:

$$F_X(x) = P(X \leq x)$$

- For a discrete random variable:
$$F_X(x) = \sum_{t \leq x} P(X = t)$$
- For a continuous random variable:
$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

3. Expectation (Mean)

The **expected value** of a random variable is its long-run average value.

- For a discrete random variable:
$$E[X] = \sum_x x \cdot P(X = x)$$
- For a continuous random variable:
$$E[X] = \int_{-\infty}^{\infty} x \cdot f_X(x) dx$$

Example: A fair six-sided die, $X = \{1, 2, 3, 4, 5, 6\}$, $P(X = x) = \frac{1}{6}$:

$$E[X] = \sum_{x=1}^{6} x \cdot \frac{1}{6} = \frac{1+2+3+4+5+6}{6} = 3.5$$

4. Variance

The **variance** measures how spread out the values of X are around the mean:

$$\text{Var}(X) = E[(X - E[X])^2] = E[X^2] - (E[X])^2$$

Example: For the six-sided die:

$$E[X^2] = \sum_{x=1}^{6} x^2 \cdot \frac{1}{6} = \frac{1^2+2^2+3^2+4^2+5^2+6^2}{6} = 15.17$$

$$\text{Var}(X) = E[X^2] - (E[X])^2 = 15.17 - (3.5)^2 = 2.92$$

5. Standard Deviation

The square root of the variance:

$$\text{Std}(X) = \sqrt{\text{Var}(X)}$$

Joint Distributions for Multiple Random Variables

When dealing with two or more random variables:

1. **Joint PMF/PDF**:
   - Joint PMF for discrete random variables X and Y:
   $$P(X = x, Y = y)$$
   - Joint PDF for continuous random variables:
   $$f_{X,Y}(x, y)$$
2. **Marginal Distribution**:
   - For a discrete random variable X:
   $$P(X = x) = \sum_y P(X = x, Y = y)$$
   - For a continuous random variable X:
   $$f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dy$$
3. **Independence**:
   - X and Y are independent if:
   $$P(X = x, Y = y) = P(X = x) \cdot P(Y = y)$$