Exercise Walkthrough: Mean and Variance of Common Discrete Distributions

Justin Lanfermann

25. June of 2025

Abstract

This document provides a detailed walkthrough for an exercise on calculating the expectation and variance of two fundamental discrete probability distributions: the Poisson and the Binomial distribution. We will leverage key definitions and properties from the script, such as the definition of expectation, the linearity of expectation, and the computational formula for variance, explaining each step in detail.

1 Overview of the Exercise

We are tasked with calculating the mean (expectation) and variance for two random variables (RVs):

(i) A discrete RV X following a Poisson distribution.

(ii) A discrete RV X following a Binomial distribution.

For each part, we will first recall the definition of the distribution's probability mass function (pmf) and then apply the definitions of expectation and variance.

2 Part (i): The Poisson Distribution

Let X be a random variable following a Poisson distribution with parameter $\lambda > 0$, denoted $X \sim \mathrm{Poi}(\lambda)$.

2.1 Recalling the Definition

From the script (Example 1.36 (vii), p. 21), the probability mass function (pmf) for a Poisson distribution is given by:

$$p(k) = P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!} \quad \text{for } k \in \{0, 1, 2, \dots\}$$

Our goal is to compute $\mathbb{E}[X]$ and $\mathrm{var}[X]$.

2.2 Calculating the Mean (Expectation)

The expectation of a discrete RV is defined as the sum of each possible value multiplied by its probability (Definition 2.1, p. 42).

$$\mathbb{E}[X] = \sum_{k=0}^{\infty} k \cdot p(k) = \sum_{k=0}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!}$$

Step 1: Simplify the Summation. The term for $k = 0$ is $0 \cdot p(0) = 0$, so we can start the summation from $k = 1$. This is a common first step when a factor of k is present.

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} k \cdot \frac{\lambda^k e^{-\lambda}}{k!}$$

Step 2: Cancel Terms. We can simplify the expression by canceling k with $k! = k \cdot (k-1)!$.

$$\mathbb{E}[X] = \sum_{k=1}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-1)!}$$

Step 3: Factor out Constants. The terms λ and $e^{-\lambda}$ can be factored out of the summation. We factor out one λ to match the power of λ with the factorial in the denominator.

$$\mathbb{E}[X] = \lambda e^{-\lambda} \sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!}$$

Step 4: Re-index the Summation. Let $m = k - 1$. As k goes from 1 to ∞, m goes from 0 to ∞. This substitution makes the sum look like a known series.

$$\mathbb{E}[X] = \lambda e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!}$$

Step 5: Recognize the Taylor Series. The sum is the Taylor series expansion of e^{λ} [1].

$$\sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = e^{\lambda}$$

Substituting this back, we get our final result for the mean:

$$\mathbb{E}[X] = \lambda e^{-\lambda} \cdot e^{\lambda} = \lambda$$

2.3 Calculating the Variance

We use the computational formula for variance (Remark 2.6, p. 45):

$$\mathrm{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$$

We already know $\mathbb{E}[X] = \lambda$. The tricky part is calculating $\mathbb{E}[X^2]$. A common technique is to first calculate $\mathbb{E}[X(X-1)]$, the second factorial moment [2], as it simplifies nicely with the $k!$ term.

Step 1: Calculate $\mathbb{E}[X(X-1)]$.

$$\mathbb{E}[X(X-1)] = \sum_{k=0}^{\infty} k(k-1) \cdot \frac{\lambda^k e^{-\lambda}}{k!}$$

The terms for $k = 0$ and $k = 1$ are zero, so the sum starts from $k = 2$.

$$= \sum_{k=2}^{\infty} k(k-1) \cdot \frac{\lambda^k e^{-\lambda}}{k!} = \sum_{k=2}^{\infty} \frac{\lambda^k e^{-\lambda}}{(k-2)!}$$

Factor out λ^2 and $e^{-\lambda}$:

$$= \lambda^2 e^{-\lambda} \sum_{k=2}^{\infty} \frac{\lambda^{k-2}}{(k-2)!}$$

Re-index with $m = k - 2$:

$$= \lambda^2 e^{-\lambda} \sum_{m=0}^{\infty} \frac{\lambda^m}{m!} = \lambda^2 e^{-\lambda} \cdot e^{\lambda} = \lambda^2$$

Step 2: Find $\mathbb{E}[X^2]$. We know that $\mathbb{E}[X(X - 1)] = \mathbb{E}[X^2 - X]$. By linearity of expectation [3], this is $\mathbb{E}[X^2] - \mathbb{E}[X]$. So, we can write:

$$\mathbb{E}[X^2] = \mathbb{E}[X(X - 1)] + \mathbb{E}[X] = \lambda^2 + \lambda$$

Step 3: Calculate the Variance. Now substitute the pieces into the variance formula:

$$\text{var}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = (\lambda^2 + \lambda) - (\lambda)^2 = \lambda$$

Conclusion for Part (i): For a random variable $X \sim \text{Poi}(\lambda)$, both its mean and variance are equal to λ.

$$\mathbb{E}[X] = \lambda \quad \text{and} \quad \text{var}[X] = \lambda$$

3 Part (ii): The Binomial Distribution

Let X be a random variable following a Binomial distribution with parameters $n \in \mathbb{N}$ and $p \in (0, 1)$, denoted $X \sim \text{Bin}(n, p)$.

3.1 Recalling the Definition and Strategy

A Binomial RV models the total number of successes in n independent and identically distributed (i.i.d.) Bernoulli trials. The pmf is given by (Example 1.36 (iv), p. 19):

$$p(k) = P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k} \quad \text{for } k \in \{0, 1, \ldots, n\}$$

While we could compute the expectation and variance using this pmf and combinatorial sums, the provided hint suggests a much more elegant approach. We can represent X as a sum of simpler random variables.

Let $Y_1, Y_2, \ldots, Y_n$ be n i.i.d. random variables, where each Y_i represents the outcome of the i-th trial. Each Y_i follows a Bernoulli distribution, $Y_i \sim \text{Ber}(p)$, where $Y_i = 1$ for a success (with probability p) and $Y_i = 0$ for a failure (with probability $1 - p$).

The total number of successes, X, is then simply the sum:

$$X = \sum_{i=1}^{n} Y_i$$

3.2 Mean and Variance of a Bernoulli RV

First, let's find the mean and variance of a single Bernoulli trial $Y \sim \text{Ber}(p)$.

$$\mathbb{E}[Y] = (1 \cdot P(Y = 1)) + (0 \cdot P(Y = 0)) = 1 \cdot p + 0 \cdot (1 - p) = p$$
$$\mathbb{E}[Y^2] = (1^2 \cdot P(Y = 1)) + (0^2 \cdot P(Y = 0)) = 1 \cdot p + 0 \cdot (1 - p) = p$$
$$\text{var}[Y] = \mathbb{E}[Y^2] - (\mathbb{E}[Y])^2 = p - p^2 = p(1 - p)$$

3.3 Calculating the Mean of X

Using the representation $X = \sum_{i=1}^{n} Y_i$, we can compute the expectation of X.

Step 1: Apply Linearity of Expectation. The expectation of a sum of random variables is the sum of their expectations. This powerful property, known as Linearity of Expectation (Proposition 2.4, p. 44), holds whether the variables are independent or not [3].

$$\mathbb{E}[X] = \mathbb{E}\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} \mathbb{E}[Y_i]$$

Step 2: Substitute the Bernoulli Mean. Since each Y_i has the same distribution, $\mathbb{E}[Y_i] = p$ for all i.

$$\mathbb{E}[X] = \sum_{i=1}^{n} p = n \cdot p$$

3.4 Calculating the Variance of X

Now we compute the variance of X.

Step 1: Apply the Variance of a Sum Formula. The variance of a sum of random variables is the sum of their variances **if they are independent** (or at least pairwise uncorrelated) (Proposition 2.8 (iv), p. 46). The problem statement specifies that the trials are i.i.d., so the Y_i are indeed independent [4].

$$\mathrm{var}[X] = \mathrm{var}\left[\sum_{i=1}^{n} Y_i\right] = \sum_{i=1}^{n} \mathrm{var}[Y_i]$$

Step 2: Substitute the Bernoulli Variance. Since each Y_i is identically distributed, $\mathrm{var}[Y_i] = p(1-p)$ for all i.

$$\mathrm{var}[X] = \sum_{i=1}^{n} p(1-p) = n \cdot p(1-p)$$

Conclusion for Part (ii): For a random variable $X \sim \mathrm{Bin}(n, p)$, the mean is np and the variance is $np(1-p)$.

$$\mathbb{E}[X] = np \quad \text{and} \quad \mathrm{var}[X] = np(1-p)$$

4 Summary

This exercise demonstrates two powerful techniques for calculating properties of distributions.

- For the **Poisson distribution**, we performed a direct calculation using its pmf, which required simplifying sums by re-indexing and recognizing the Taylor series for e^λ.

- For the **Binomial distribution**, we used a more elegant modeling approach by representing the RV as a sum of i.i.d. Bernoulli RVs. This allowed us to use the powerful properties of linearity of expectation and variance of sums of independent variables to arrive at the solution much more easily than by direct computation.

In-depth Explanations

[1] Taylor Series for e^λ: From calculus, we know that the exponential function e^x can be represented by an infinite sum, its Taylor (or Maclaurin) series, which converges for all real (and complex) x:

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots$$

This identity is fundamental in many areas of mathematics and is particularly useful in probability theory for working with distributions like Poisson.

[2] Factorial Moments: The quantity $\mathbb{E}[X(X-1)\dots(X-k+1)]$ is known as the k-th factorial moment of X. For discrete distributions whose pmfs involve a factorial in the denominator (like Poisson), calculating factorial moments is often simpler than calculating ordinary moments ($\mathbb{E}[X^k]$) directly. This is because the term $k(k-1)\dots$ in the sum cancels neatly with the $k!$ in the pmf. From the factorial moments, one can then recover the ordinary moments.

[3] Linearity of Expectation: For any two random variables X and Y defined on the same probability space, and any constants $a, b \in \mathbb{R}$, the expectation operator is linear:

$$\mathbb{E}[aX + bY] = a\mathbb{E}[X] + b\mathbb{E}[Y]$$

This property is incredibly powerful because it holds **regardless of whether X and Y are independent**. This is a cornerstone of probability theory and simplifies many calculations.

[4] Variance of a Sum of Independent Variables: For any two random variables X and Y, the variance of their sum is given by:

$$\text{var}[X + Y] = \text{var}[X] + \text{var}[Y] + 2\text{cov}[X, Y]$$

where $\text{cov}[X, Y]$ is the covariance between X and Y. A key property is that if X and Y are independent, their covariance is zero. Therefore, for independent random variables, the formula simplifies to:

$$\text{var}[X + Y] = \text{var}[X] + \text{var}[Y]$$

This extends to a sum of any number of pairwise independent random variables. It is crucial to remember the independence requirement, which distinguishes this from the more general linearity of expectation.