Mathematical Statistics I

Chapter 4: Expected Values

Jesse Wheeler

Contents

1 Discrete random variables

Introduction

- This material comes primarily from Rice (2007, Chapter 4), though I'm going to deviate slightly.
- We will cover the ideas of expected value, variance, as well has higher-order moments.
- This includes topics such as conditional expectation, which is one of the fundamental ideas behind many branches of statistics and machine learning.
- For instance, most regression / prediction algorithms are built with the idea of minimizing some conditional expectation.

Expectation: Discrete random variables

- We will begin by defining the expectation for discrete random variables.
- I'm going to deviate from our textbook in this definition.

Definition: Expectation of discrete random variables
Let X be a discrete random variable with pmf $p(x)$, which takes values in the space $\mathcal{X}$. The *expected value* of $g(X)$ is

$$E\big(g(X)\big) = \sum_{x \in \mathcal{X}} g(x)\, p(x).$$

In particular, for $g(x) = x$, we have

$$E(X) = \sum_{x \in \mathcal{X}} x\, p(x).$$

- This is not the most mathematically precise definition of expectation, but a more complete treatment of the topic is outside the scope of this course (See Resnick, 2019).
- The definition is only applicable if the sum is finite.
- The concept of the expected value parallels the notion of a *weighted average*.

- That is, we weight each possibility $x \in \mathcal{X}$ by their corresponding probability: $\sum_x x\, p(x)$.
- $E(X)$ is also referred to as the *mean* of X, and is typically denoted μ or μ_X.
- If the function p is thought of as a weight, then $E(X)$ is the center; that is, if we place the mass $p(x_i)$ at the points x_i, then the balancing point is $E(X)$.
- Like with the pmf and cdf, we often use subscripts to denote which probability law we are using for the expectation, it if is not clear: $E_X(X)$.

Roulette
A roulette wheel has the numbers 1 through 36, as well as 0 and 00. If you bet \$1 that an odd number comes up, you win or lose \$1 according to whether that event occurs. If X denotes your net gain, $X = 1$ with probability 18/38 and $X = -1$ with probability 20/28. The expected value of X is

$$E(X) = 1 \times \frac{18}{38} + (-1) \times \frac{20}{38} = -\frac{1}{19}.$$

- As you might imagine, the expected value coincides in the limit with the actual average loss per game, if you play many games (Chapter 5).
- Most casino games have a negative expected value by design; you may win some money, but if a large number of games are played, the house will come out on top.

Geometric Random Variable
Suppose that items are produced in a plant are independently defective with probability p. If items are inspected one by one until a defective item is found, then how many items must be inspected on average?
Let X denote the number of items inspected, up-to and including the first defective item. X is geometrically distributed, which as pmf

$$p(k) = P(X = k) = p\,(1-p)^{k-1}.$$

Therefore

$$\begin{aligned} E(X) &= \sum_{k=1}^{\infty} kp\,(1-p)^{k-1} \\ &= p \sum_{k=1}^{\infty} k\,(1-p)^{k-1}. \end{aligned}$$

To work out this summation, we will use a trick that is sometimes useful for infinite series. First, lets define $q = 1 - p$, and note that $0 < q < 1$. Then, the sum becomes

$$E(X) = p \sum_{k=1}^{\infty} k\, q^{k-1}.$$

You might notice that the summand is a power-rule derivative:

$$\frac{d}{dq} q^k = k\, q^{k-1}.$$

This fact is going to be useful, because the left-hand side of this derivative equation is a geometric sum, which we know how to calculate:

$$\sum_{k=1}^{\infty} q^k = \sum_{k=1}^{\infty} q\, q^{k-1} = q \sum_{j=0}^{\infty} q^j = \frac{q}{1-q}.$$

Thus, what we would like to do is write

$$\frac{d}{dq}\Big(\frac{q}{1-q}\Big) = \frac{d}{dq}\Big(\sum_{k=1}^{\infty} q^k\Big) \stackrel{?}{=} \sum_{k=1}^{\infty} \frac{d}{dq} q^k = \sum_{k=1}^{\infty} k q^{k-1}.$$

Now we can easily calculate the left-hand side to be $\frac{1}{(1-q)^2}$, and therefore we want to make the conclusion

$$\sum_{k=1}^{\infty} k\, q^{k-1} \stackrel{?}{=} \frac{d}{dq}\Big(\frac{q}{1-q}\Big) = \frac{1}{(1-q)^2}.$$

The question is: **Can we move the derivative inside of the infinite sum?** For this particular case, the answer is *yes*. In more advanced analysis classes, you learn methods for justifying this step rigorously using uniform convergence. Specifically, what we need is for uniform convergence of the partial sums and their derivatives. Fortunately for this class, all of the sums (and integrals) we will consider will be "well-behaved" and will satisfy these conditions.
With this sorted out, we can now use our trick to finish the calculation:

$$\begin{aligned} E(X) &= p\sum_{k=1}^{\infty} k\,(q)^{k-1} \\ &= p\frac{1}{(1-q)^2} \\ &= \frac{p}{p^2} = \frac{1}{p}. \end{aligned}$$

Poisson Distribution
The Poisson(λ) distribution has pmf $p(k) = \frac{\lambda^k}{k!}e^{-\lambda}$, for all $k \geq 0$. Thus, if $X \sim \text{Pois}(\lambda)$, then what is $E[X]$?

$$\begin{aligned} E[X] &= \sum_{k=0}^{\infty} \frac{k\,\lambda^k}{k!} e^{-\lambda} \\ &= e^{-\lambda}\sum_{k=0}^{\infty} \frac{k\,\lambda^{k-1}\cdot\lambda}{k!} \\ &= \lambda e^{-\lambda}\sum_{k=1}^{\infty} \frac{\lambda^{k-1}}{(k-1)!} \\ &= \lambda e^{-\lambda}\sum_{j=0}^{\infty} \frac{\lambda^j}{j!} \\ &= \lambda e^{-\lambda} e^{\lambda} = \lambda. \end{aligned}$$

2 Continuous random variables

Expectation: Continuous random variables

Acknowledgments

- Compiled on August 22, 2025 using R version 4.5.1.

- We acknowledge students and instructors for previous versions of this course / slides.

References

Resnick S (2019). *A probability path*. Springer. 2

Rice JA (2007). *Mathematical statistics and data analysis*, volume 371. 3 edition. Thomson/Brooks/Cole Belmont, CA. 1