STA 111 (Summer Session I)

Lecture Four – Expectation and Continuous Random Variables

Instructor: Olanrewaju (Michael) Akande

Department of Statistical Science, Duke University

May 23, 2016

Outline

- Questions from Last Lecture.
- Expectation and Variance
- Continuous Random Variables
- The Normal Distribution
- Recap

Introduction

– In the last lecture we learned about random variables.

– We also talked about some important discrete distributions.

– Today we will continue with how to calculate expectations for random variables.

– Next we will revisit the definition of continuous random variables and talk about the most important continuous distribution.

Expectation

– The expected value of a random variable is the average value of an infinite number of draws. It is often denoted by μ. Mathematically,

$$\mathbb{E}[X] \equiv \mu \equiv \sum_{\text{all possible } X} xp(x).$$

– This is a weighted average of the possible values of the random variable, where the weights are the probabilities that a value will occur.

– For example, the expected value of a fair die is 3.5. Why?

Expectation (Cont'd)

– The expected value of a function of a random variable, say $h(X)$, is the long-run average value of repeated evaluations of that function. Or

$$\mathbb{E}[h(X)] = \sum_{\text{all possible } X} h(x)p(x).$$

– For example, the expected value of X^2 for a fair die is

$$\sum_{i=1}^{6} i^2 * \frac{1}{6} = 15.1667$$

– Some functions are especially interesting. For example, the linear function $aX + b$ has the useful property that

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

Expectation (Cont'd)

– The variance of a random variable is the expected value of $h(X) = (X - \mu)^2$ where μ is the expected value of X.

– The variance is often written as σ^2, and its units are the squares of the original units. It provides a measure of how spread out the sample is, since it is the average squared distance of an observation from μ. A little algebra shows:

$$\begin{aligned}
\text{Var}[X] &= \mathbb{E}[(X-\mu)^2] \\
&= \sum_{\text{all } X} (x-\mu)^2 p(x) \\
&= \sum_{\text{all } X} (x^2 - 2x\mu + \mu^2) p(x) \\
&= \mathbb{E}[X^2] - 2\mu\mathbb{E}[X] + \mu^2 \\
&= \mathbb{E}[X^2] - \mu^2
\end{aligned}$$

$$\text{Var}[aX+b] = \mathbb{V}[aX+b] = a^2\mathbb{V}[X].$$

Expectation (Cont'd)

– To measure how spread out a distribution is, we mostly use the standard deviation (or σ). This is the square root of the variance.

– The variance of the result of a roll of a fair die is $\mathbb{E}[X^2] - \mathbb{E}[X]^2 = 15.1667 - (3.5)^2 = 2.9167$. Its standard deviation is $\sqrt{2.9167} = 1.7078$.

– One can show that the mean, variance and standard deviation of the $\text{Bin}(n, p)$ distribution are np, $np(1-p)$ and $\sqrt{np(1-p)}$, respectively.

– For the $\text{Pois}(\lambda)$ distribution, the mean is λ, the variance is λ, and the standard deviation is $\sqrt{\lambda}$. We will do this in class!

Expectation (Cont'd)

How do we know that the mean of the binomial is np?

$$
\begin{aligned}
\mathbb{E}[X] &= \sum_{\text{all } x} x p(x) \\
&= \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} x \frac{n!}{x!(n-x)!} p^x (1-p)^{n-x} \\
&= \sum_{x=0}^{n} \frac{n!}{(x-1)!(n-x)!} p^x (1-p)^{n-x}.
\end{aligned}
$$

Expectation (Cont'd)

Make the change of variable $y = x - 1$ to get

$$\begin{aligned}\mathbb{E}[X] &= np \sum_{y=0}^{n-1} \frac{(n-1)!}{y!(n-1-y)!} p^y (1-p)^{n-1-y} \\ &= np \sum_{y=0}^{n-1} x \binom{n-1}{y} p^y (1-p)^{n-1-y}.\end{aligned}$$

We recognize the summation as the sum of binomial probabilities for a $\text{Bin}(n-1, p)$ distribution. The sum of all possible probabilities is always 1, so the mean is np.

A similar argument works for $\mathbb{E}[X^2]$, from which one can verify that the variance is $np(1-p)$. (Recall that the variance is $\mathbb{E}[X^2] - (\mathbb{E}[X])^2$.) We will do this in class!

Continuous Random Variables

– Recall that a random variable X has a continuous distribution if X can take on infinite values. For continuous distributions, specific values have probability zero. One can only assign probabilities to intervals

– For example, there is a positive probability of tossing exactly five heads in 10 flips. But there is zero probability of finding someone who is exactly five feet tall, when height is measured to an infinite number of decimal places.

– To define probabilities for intervals, we use a density function. A density function is any function $f(x)$ such that

- $f(x)$ is non-negative
- $f(x)$ integrates to 1.

Then

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f(x)\, dx.$$

Continuous Random Variables (Cont'd)

This definition of the density function ensures the probability axioms are satisfied:

1. All probabilities are between 0 and 1, inclusive,
2. $\mathbb{P}[-\infty < X < \infty] = 1$, and
3. If A and B are disjoint intervals, then

$$\mathbb{P}[X \in A \text{ or } X \in B] = \mathbb{P}[X \in A] + \mathbb{P}[X \in B].$$

It turns out to be useful to define the cumulative distribution function (or cdf) as

$$F(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^{x} f(y)\, dy.$$

Then

$$\mathbb{P}[a \leq X \leq b] = \int_{a}^{b} f(x)\, dx = F(b) - F(a).$$

Continuous Random Variables (Cont'd)

The expected value of a continuous random variable is

$$\mathbb{E}[X] \equiv \mu \equiv \int_{-\infty}^{\infty} xf(x)\,dx.$$

Clearly, this is similar to the definition of expected value for a discrete random variable:

$$\mathbb{E}[X] \equiv \mu \equiv \sum_{\text{all } X} xp(x).$$

Analogously, the expected value of a function $h(X)$ of a continuous random variable is

$$\mathbb{E}[h(X)] \equiv \int_{-\infty}^{\infty} h(x)f(x)\,dx.$$

As before, $\mathbb{E}[aX + b] = a\mathbb{E}[X] + b$. And $\mathbb{V}[X] = \mathbb{E}[X^2] - (\mathbb{E}[X])^2$.

The Normal Distribution

The most famous continuous distribution is the normal distribution (also called the Gaussian distribution or the bell-shaped curve). It's p.d.f.s cannot be integrated in closed form, and hence tables of the c.d.f. or computer programs are necessary in order to compute probabilities and quantiles.

The distribution was named after Carl Friedrich Gauss, the greatest mathematician in history. He proved the fundamental theorem of algebra four ways, inventing a new branch of mathematics each time. He worked in number theory, co-invented the telegraph, and discovered non-Euclidean geometry, but did not publish, fearing controversy.

The Normal Distribution (Cont'd)

People believe the normal distribution describes IQ, height, rainfall, measurement error, and many other features. This is only approximately true. However, the random variables studied in various physical experiments often have distributions that are approximately normal.

A normal distribution with mean μ and standard deviation σ has the p.d.f:

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]$$

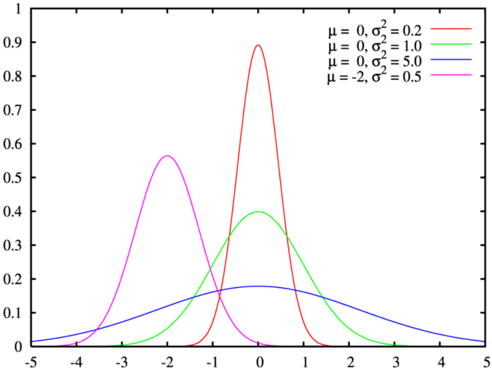
for $-\infty < \mu < \infty$ and $\sigma \geq 0$.

The μ is the mean of the entire population, whereas $\bar{X}$ is used to denote the mean of a sample from the population. Similarly, σ is the standard deviation of the entire population, whereas *sd* is used to denote the standard deviation of a sample.

The standard normal has $\mu = 0$ and $\sigma = 1$.

The Normal Distribution (Cont'd)

The mean of a normal distribution shows where it is centered. The standard deviation of a normal distribution shows how spread out the normal is.



The Standard Normal Distribution

We will learn more about the distribution in the next lecture, but for now lets practice reading the standard normal table. You should also practice using it at home.

What proportion of a standard normal population has values between -1.5 and 1.5? From the table, the proportion is 1 - 2 * 0.067 = 0.866 (i.e., the total area under the curve is one, and we subtract the upper tail area and the symmetric lower tail area).

Now go the other way. About 80% of the population lies between what two values that are centered at 0? The answer is about -1.28 and 1.28 (the table shows that 10% of the area is above 1.28, and symmetrically, 10% is below -1.28).

The Standard Normal Distribution

Some problems to think about:

- What is the value of z such that 25% of a standard normal population is larger than that value? (Ans: about 0.67)
- What is the value for which about 90% of the population is smaller? (Ans: about 1.28)
- What proportion of the population has a value larger than -1? (Ans: about 0.841)
- What proportion of the population has a value less than -0.3? (Ans: about 0.382)

In this class, use the nearest value in the table. Do not use a number from your calculator!

Recap

We discussed the following:

- Calculating expectation and variance
- The normal distribution