Expectation and Other Parameters

Expectation (denoted $E[X], \mu_X$, or μ) – For a random variable *X*, the expectation of *X* (aka expected value of *X*, or mean of *X*) is the weighted average of the values of *supp(X)*. The weights are the corresponding values of the pdf.

For a discrete random variable we have

$$E[X] = \sum_{x \in \sup p(X)} x \cdot p(x) = x_1 \cdot p(x_1) + x_2 \cdot p(x_2) + \cdots$$

For a continuous random variable we have

$$E[X] = \int_{-\infty}^{\infty} x \cdot f(x)dx .$$

If *X* is non-negative, it can be shown that $E[X] = \int_0^{\infty} x \cdot f(x)dx = \int_0^{\infty} S(x)dx.$.

If *h* is a function of the random variable *X*, then the expectation of *h(X)* is

i) If *X* is discrete, $E[h(X)] = \sum_{x \in \sup p(X)} h(x) \cdot p(x)$

ii) If *X* is continuous, $E[h(X)] = \int_{-\infty}^{\infty} h(x) \cdot f(x)dx$.

Notice that the expectation formulas above are a special case of these formulas with *h(X) = X*.

Example: Suppose *X* is a discrete random variable with *p(0) = 0.5, p(1) = 0.2,* and *p(4) = 0.3*. Find $E[X^{0.5}]$, and separately find $E[3X + 2]$.

Let's again draw a probability distribution table.

X	$X^{0.5}$	$3X + 2$	$p(x)$
0	0	2	0.5
1	1	5	0.2
4	2	14	0.3

Then $E[X^{0.5}] = E[\sqrt{X}] = 0 \cdot 0.5 + 1 \cdot 0.2 + 2 \cdot 0.3 = 0.8$ and $E[3X + 2] = 2 \cdot 0.5 + 5 \cdot 0.2 + 14 \cdot 0.3 = 6.2$. Notice that since $E[X] = 0 + 0.2 + 1.2 = 1.4$, then $E[X^{0.5}] \neq (E[X])^{0.5}$. However we do have that $E[3X + 2] = 3E[X] + 2$. In general, we have the formula $E[a \cdot g(X) + b \cdot h(X) + c] = a \cdot E[g(X)] + b \cdot E[h(X)] + c$ where *a, b,* and *c* are constants.

Other Distribution Parameters and Relationships Among Them

The n^{th} moment of the random variable X is defined to be $E[X^n]$. If the mean of X is $\mu_X = \mu$, then the n^{th} central moment of X about the mean μ is $E[(X-\mu)^n]$.

The variance of the random variable X is denoted by $Var(X)$, $V(X)$, σ_X^2, or σ^2, and is defined to be the 2$^{\text{nd}}$ central moment of X about the mean μ.
We have

$$Var(X) = E[(X-\mu)^2] = E[X^2 - 2\mu X + \mu^2] = E[X^2] - 2\mu E[X] + \mu^2 = E[X^2] - (E[X])^2$$

An important property of the variance is that if $Y = aX + b$, where a and b are constants, then

$$Var(Y) = Var(aX+b) = a^2 Var(X).$$

The standard deviation of the random variable X is the square root of the variance and is thus denoted by σ_X or σ. In symbols,

$$\sigma_X = \sqrt{Var(X)}$$

The coefficient of variation of the random variable X is the ratio of the standard deviation of X to the mean of X. In symbols,

$$CV_X = \frac{\sigma_X}{\mu_X}$$

We use examples to illustrate the concept of percentiles of a distribution.

The median of a distribution is the 50$^{\text{th}}$ percentile of the distribution.

The mode of a distribution is any value of the random variable X at which the pdf is maximized.

The moment generating function (mgf) of the random variable X is denoted $M_X(t)$, $m_X(t)$, $M(t)$, or $m(t)$ and is defined to be $M_X(t) = E[e^{tX}]$.

Properties of mgf's:

1. $M_X(0) = 1$

2. If X_1 and X_2 are random variables and $M_{X_1}(t) = M_{X_2}(t)$, then $X_1 \sim X_2$.

3. $E[X] = \frac{d}{dt} M_X(t)|_{t=0} = M'_X(0)$, and in general $E[X^n] = M_X^{(n)}(0)$.

4. If we define $R_X(t) = \ln(M_X(t))$, then

$$R'_X(0) = \frac{d}{dt} R_X(t)|_{t=0} = \frac{M'_X(t)}{M_X(t)}|_{t=0} = \frac{M'_X(0)}{M_X(0)} = \frac{E[X]}{1} = E[X],$$ and similarly

$$R''_X(0) = Var(X)$$

Since the mgf is defined to be the expectation of the random variable e^{tX}, it is useful to recall the Taylor series expansion for the exponential function:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} = 1 + x + \frac{x^2}{2} + \frac{x^3}{6} + \cdots$$

The skewness of a distribution is defined to be $\frac{E[(X-\mu)^3]}{\sigma^3}$. If this value is negative then we say the random variable is skewed to the left (of its mean) and if it is positive then we say the random variable is skewed to the right (of its mean).

Chebyshev's Inequality: If X is a random variable with mean μ and standard deviation σ then for any real number $r > 0$ we have

$$\Pr(|X - \mu| > r \cdot \sigma) \leq \frac{1}{r^2}$$

A picture is worth a thousand words: