Math 302.102 Fall 2010
Summary of Lecture from October 27, 2010

Suppose that X is a continuous random variable with density f. In applications, X might represent (i) the payout of a bet on a chance experiment, (ii) the lifetime of a manufactured component of a physical system such as a lightbulb, circuit board, battery, crystal oscillator, or catalytic converter, (iii) the value of a given stock at some fixed time in the future, (iv) a temperature, pressure, displacement, mass, or some other measurable characteristic of a physical body subject to random interactions with the surrounding medium, or (v) etc. We already know how to compute probabilities associated with X, namely

$$\mathbf{P}\left\{X \in B\right\} = \int_B f(x)\,\mathrm{d}x$$

for any $B \subseteq \mathbb{R}$.

The moments of a random variable are a useful way to *summarize* that random variable. By definition, the kth *moment* of X is defined as

$$\mathbb{E}(X^k) = \int_{-\infty}^{\infty} x^k f(x)\,\mathrm{d}x$$

for $k = 1, 2, 3, \ldots$.

The first moment, $\mathbb{E}(X) = \mathbb{E}(X^1)$, is also called the *expected value* (or *expectation* or *mean*) of X. It is the continuous analogue of a weighted average. That is, since the density function of X can be interpreted as the infinitesmal likelihood of a particular value of X, the "continuous weighted average" of X is simply

$$\int_{-\infty}^{\infty} x f(x)\,\mathrm{d}x = \mathbb{E}(X).$$

Perhaps it is better to just think of $\mathbb{E}(X)$ as the *average* payout of a bet, or the *average* lifetime of a manufactured component, or the *average* etc., when X has density f.

Remark. Since the expected value of X is defined as a definite integral, it is just a number. As such, we know that $\mathbf{P}\left\{X = \mathbb{E}(X)\right\} = 0$ (since the probability that a continuous random variable takes on any fixed value is 0). However, we will soon learn a way to estimate the *deviations* of X from $\mathbb{E}(X)$.

The use of the word *moment* in probability is taken (loosely) from physics where it refers to many different concepts including moment of force or moment of inertia. It has also been used historically in mathematics to describe the "shape" of a set of points.

The importance of the higher order moments ($k = 2, 3, \ldots$) is clear when we consider the kth *central moment* of X denoted by μ_k and defined as follows. Set $\mu_1 = \mu = \mathbb{E}(X)$ and for $k = 2, 3, \ldots$ define

$$\mu_k = \mathbb{E}\left([X - \mathbb{E}(X)]^k\right) = \mathbb{E}\left([X - \mu]^k\right) = \int_{-\infty}^{\infty} (x - \mu)^k f(x)\,\mathrm{d}x.$$

The second central moment is called the *variance* and it a measure of how much a random variable deviates from its expected value.

Often, we write

$$\sigma^2 = \mathrm{Var}(X) = \mu_2$$

for the variance. The square root of the variance is called the *standard deviation.* We sometimes write

$$\sigma = \mathrm{SD}(X) = \sqrt{\mathrm{Var}(X)}$$

for the standard deviation. (In physics, the second central moment is called the moment of inertia about the centre of mass.)

The normalized third central moment is called the *skewness* and is defined by

$$\gamma_1 = \frac{\mu_3}{\sigma^3} = \frac{\mu_3}{\mu_2^{3/2}}.$$

It is a measure of asymmetry (or skewness) of the density/distribution of the random variable.

The normalized fourth central moment is called the *kurtosis* and is defined by

$$\gamma_2 = \frac{\mu_4}{\sigma^4} = \frac{\mu_4}{\mu_2^2}.$$

It is a measure of the peakedness of the density/distribution of the random variable.

Remark. None of the other higher (normalized/central) moments are given special names. For Math 302, we will be concerned almost entirely with the expected value (i.e., the first moment or mean) and the variance (i.e., the second central moment). The primary uses of the skewness and kurtosis are in statistics in the context of parametric tests. The *method of moments* is a somewhat useful technique for estimating parameters in a statistical model whereby a system of equations involving the population moments and the sample moments is solved for the unknown parameters.

The *moment generating function* of X is the function $m : \mathbb{R} \to \mathbb{R}$ defined by

$$m(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)\,\mathrm{d}x.$$

It has the property that the kth moment of X can be determined by evaluating the kth derivative of the moment generating function at $t = 0$. That is,

$$m^{(k)}(0) = \frac{\mathrm{d}^k}{\mathrm{d}t^k} m(t)\bigg|_{t=0} = \mathbb{E}(X^k).$$

The proof proceeds as follows. By interchanging the derivative and the integral, we find

$$\begin{aligned}\frac{\mathrm{d}^k}{\mathrm{d}t^k} m(t) = \frac{\mathrm{d}^k}{\mathrm{d}t^k} \int_{-\infty}^{\infty} e^{tx} f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} \left[\frac{\mathrm{d}^k}{\mathrm{d}t^k} e^{tx} f(x)\right] \mathrm{d}x &= \int_{-\infty}^{\infty} f(x) \left[\frac{\mathrm{d}^k}{\mathrm{d}t^k} e^{tx}\right] \mathrm{d}x \\ &= \int_{-\infty}^{\infty} x^k e^{tx} f(x)\,\mathrm{d}x.\end{aligned}$$

Thus, evaluating at $t = 0$ gives

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} m(t)\bigg|_{t=0} = \int_{-\infty}^{\infty} x^k e^{0\cdot x} f(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} x^k f(x)\,\mathrm{d}x = \mathbb{E}(X^k)$$

as required.