Moments and the moment generating function
Math 217 Probability and Statistics
Prof. D. Joyce, Fall 2014

Today we'll prove the central limit theorem.

A proof of the central limit theorem by means of moment generating functions. This theorem says that for any distribution X with a finite mean μ and variance σ^2, the sample sum S and also the sample mean $\overline{X}$ approach a normal distribution. More specifically, the standardized sample sum S^* and the standardized sample mean $\overline{X}^*$, which are, in fact, the same thing, approach the standard normal distribution.

Consider a distribution X with mean μ and variance σ^2. Let $X_1, \ldots, X_n$ be a sample from this distribution with sample sum $S_n = X_1 + \cdots + X_n$. We'll show that the standardized sum

$$S_n^* = \frac{S_n - n\mu}{\sigma\sqrt{n}}$$

approaches a standard normal distribution by showing that its moment generating function $m_{S_n^*}(t)$ approaches the moment generating function for the standard normal distribution Z, which is $m_Z(t) = e^{t^2/2}$.

Last time we saw that the standardized sum S_n^* has the generating function

$$m_{S_n^*}(t) = e^{-\sqrt{n}\mu t/\sigma} \left(m_X(\tfrac{t}{\sigma\sqrt{n}}) \right)^n$$

First, we can simplify things a little bit by assuming that the mean μ is 0. That is, if $Y = X - \mu$, then the mean of Y is 0, Y has the same standard deviation σ as X, and X and Y have the same standardized sum. Thus, if the central limit theorem holds for distributions Y with mean 0, then it holds for distributions with any mean μ.

Thus, we assume the mean of X is $\mu = 0$ and continue with the proof. Likewise, we can simplify things a little bit more by assuming that the standard deviation σ is 1. That is, if $Y = X/\sigma$, then the mean of Y is 0, its standard deviation is 1, and X and this Y have the same standardized sum. Thus, if the theorem holds for distributions Y with mean 0 and standard deviation 1, then it holds for any distribution with mean μ and standard deviation σ.

Thus, we assume the standard deviation of X is $\sigma = 1$ and continue with the proof.

Now, with $\mu = 0$ and $\sigma = 1$, we have

$$m_{S_n^*}(t) = \left(m_X\left(\frac{t}{\sqrt{n}}\right)\right)^n.$$

Note that since $\mu = 0$, the second moment μ_2 is just the variance $\sigma^2 = 1$ because in any case $\sigma^2 = E(X^2) - \mu^2 = \mu_2 - \mu^2$. Thus $\mu_2 = 1$.

The idea of the proof is to use just the first couple of terms of the power series for the moment generating function $m_X(t)$ because the rest of the terms are very small. More precisely, the moment generating function for X can be written as

$$\begin{aligned} m_X(t) &= 1 + \frac{1}{2}\mu_2 t^2 + Rt^3 \\ &= 1 + \frac{1}{2}t^2 + Rt^3 \end{aligned}$$

where Rt^3 includes the remainder of the power series, that is, all the higher terms for the power series. (The existence of this power series assumes that all higher moments exist. That assumption is not necessary because even without a power series, Taylor's theorem givesa remainder term that does the same job.)

In the expression for the generating function $m_{S_n^*}(t)$ of the standardized sample sum, we have the expression $m_X\left(\frac{t}{\sqrt{n}}\right)$, so we'll analyze it next.

$$\begin{aligned} m_X\left(\frac{t}{\sqrt{n}}\right) &= 1 + \frac{1}{2}^2\left(\frac{t}{\sqrt{n}}\right)^2 + R\left(\frac{t}{\sqrt{n}}\right)^3 \\ &= 1 + \frac{t^2}{2n} + R\left(\frac{t}{\sqrt{n}}\right)^3 \end{aligned}$$

The remainder term $R\left(\frac{t}{\sqrt{n}}\right)^3$ is small compared to t^2/n in the sense that as n approaches ∞, the ratio of $R\left(\frac{t}{\sqrt{n}}\right)^3$ to t^2/n approaches 0. Changing notation slightly, we can say that

$$m_X\left(\frac{t}{\sqrt{n}}\right) = 1 + \frac{t^2}{2n} + R'$$

where R' is small compared to t^2/n.

Now we can analyze the generating function $m_{S_n^*}(t)$ of the standardized sample sum itself.

$$\begin{aligned} m_{S_n^*}(t) &= \left(m_X\left(\frac{t}{\sqrt{n}}\right)\right)^n \\ &= \left(1 + \frac{t^2}{2n} + R'\right)^n. \end{aligned}$$

We'll show that as $n \to \infty$, this last expression approaches $e^{t^2/2}$. The difficulty is that n appears both as an exponent and as a denominator. The n in the exponent suggests the limit might be ∞, but the n in the denominator suggests the limit might be 1. It turns out to be neither.

There are various ways to proceed. We'll take natural logs to get

$$\log m_{S_n}(t) = n \log\left(1 + \frac{t^2}{2n} + R'\right).$$

For small values of x, the natural $\log(1+x)$ is about x. Indeed,

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots.$$

Hence, $\log\left(1 + \frac{t^2}{2n} + R'\right)$ equals

$$\frac{t^2}{2n} + R' - \frac{\left(\frac{t^2}{2n} + R'\right)^2}{2} + \frac{\left(\frac{t^2}{2n} + R'\right)^3}{3} - \cdots.$$

Since all the terms except the first are small compared to t^2/n (in the same sense mentioned above), we can rewrite the expression as

$$\frac{t^2}{2n} + R''$$

where R'' is small compared to t^2/n. We can now say that

$$\log m_{S_n^*}(t) = \frac{t^2}{2} + nR''.$$

But R'' is small compared to t^2/n, so nR'' is also small compared to t^2. Thus, as $n \to \infty$, $\log m_{S_n^*}(t) \to t^2/2$. Finally, exponentiating, we get

$$m_{S_n^*}(t) \to e^{t^2/2}$$

as $n \to \infty$. Thus, the generation function $m_{S_n^*}(t)$ for S_n^* approaches the generating function $e^{t^2/2}$ of the standard normal distribution. Hence, S_n^* approaches the standard normal distribution. Q.E.D.

Math 217 Home Page at
http://math.clarku.edu/~djoyce/ma217/