

Stat 134 Fall 2011: a proof of the central limit theorem

Michael Lugo

October 18, 2011

Pitman does not give a proof of the central limit theorem. Using moment generating functions and some results from analysis that should be intuitively plausible, it's possible to give a proof. This proof follows the presentation of *Mathematical Statistics and Data Analysis*, John A. Rice, third edition; this is the textbook often used in Stat 135.

First we need a few facts about moment generating functions. Recall that the moment generating function of a random variable X is $M_X(t) = E(e^{tX})$. We'll write this as

$$\int_{-\infty}^{\infty} e^{tx} f(x) dx$$

and think of f as a continuous random variable, although the result holds for discrete random variables as well. If we differentiate the mgf r times, we get

$$M^{(r)}(t) = \frac{d^r}{dt^r} \int_{-\infty}^{\infty} e^{tx} f(x) dx = \int_{-\infty}^{\infty} \left(\frac{d^r}{dt^r} e^{tx} f(x) \right) dx = \int_{-\infty}^{\infty} x^r e^{tx} f(x) dx.$$

(The third equality requires some justification, in order to interchange differentiation and integration, but this should seem believable.) Letting $t = 0$ we get

$$M^{(r)}(0) = \int_{-\infty}^{\infty} x^r f(x) dx = E(X^r).$$

We need to know the moment generating function of a product, and how mgfs behave under linear transformations.

Proposition 1. *If X and Y are independent random variables with mgfs M_X and M_Y , and $Z = X + Y$, then $M_Z(t) = M_X(t)M_Y(t)$.*

Proof. The proof is a one-liner:

$$M_Z(t) = E(e^{tZ}) = E(e^{tX+tY}) = E(e^{tX}e^{tY}) = E(e^{tX})E(e^{tY}) = M_X(t)M_Y(t)$$

where the third equality follows from the assumption of independence. □

Proposition 2. *If X is a random variable with mgf M_X , and $Y = a + bX$, then $M_Y(t) = e^{at}M_X(bt)$*

Proof. $M_Y(t) = E(e^{tY}) = E(e^{at+btX}) = E(e^{at}e^{btX}) = e^{at}E(e^{btX}) = e^{at}M_X(bt)$. □

Next, we need two definitions and a theorem. The *cumulative distribution function* (cdf) of a random variable X is $F(x) = P(X \leq x)$. If X_1, X_2, \dots , are a sequence of random variables with cdfs F_1, F_2, \dots , and X is a random variable with cdf F , then X_n *converges in distribution* to X if $\lim_{n \rightarrow \infty} F_n(x) = F(x)$ at every point at which F is continuous.

(The limit can be reinterpreted as $\lim_{n \rightarrow \infty} P(X_n \leq x) = P(X \leq x)$.)

The central theorem that we'll need is the following.

Theorem 3. *Let F_n be a sequence of cumulative distribution functions with the correspond moment generating functions M_n . Let F be a cdf with the mgf M . If $M_n(t) \rightarrow M(t)$ for all t in an open interval containing zero, then $F_n(x) \rightarrow F(x)$ for all x at which F is continuous.*

This says that if the moment generating functions of a sequence of distributions approach some limit, then that limiting mgf is the mgf of the limit of that sequence of distributions. So in order to prove the CLT, it will be enough to show that the mgf of a standardized sum of n independent, identically distributed random variables approaches the mgf of a standard normal as $n \rightarrow \infty$.

Theorem 4 (Central limit theorem). *Let X_1, X_2, \dots be a sequence of independent random variables having mean μ and variance σ^2 . Let each X_i have the cdf $P(X_i \leq x) = F(x)$ and the moment generating function $M(t) = E(e^{tX_i})$. Let $S_n = \sum_{i=1}^n X_i$. Then*

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - \mu n}{\sigma\sqrt{n}} \leq x\right) = \Phi(x)$$

for $-\infty < x < \infty$.

(Pitman gives $\lim_{n \rightarrow \infty} P(a \leq S_n/(\sigma\sqrt{n}) \leq b) = \Phi(b) - \Phi(a)$ for all $-\infty < a < b < \infty$; this is equivalent, since the limit of a difference is the difference of the limits.)

Proof. It suffices to do the proof in the case $\mu = 0$. If $\mu \neq 0$, let $Y_i = X_i - \mu$ for each i . Let $T_n = Y_1 + Y_2 + \dots + Y_n$. Then we have

$$\lim_{n \rightarrow \infty} P\left(\frac{S_n - \mu n}{\sigma\sqrt{n}} \leq x\right) = P\left(\frac{T_n}{\sigma\sqrt{n}} \leq x\right)$$

and so it suffices to prove the central limit theorem in the case $\mu = 0$.

Let $Z_n = (S_n)/(\sigma\sqrt{n})$. We'll show that the mgf of Z_n tends to the mgf of the standard normal distribution. Since S_n is a sum of independent random variables,

$$M_{S_n}(t) = [M(t)]^n$$

and by Proposition 2, we have

$$M_{Z_n}(t) = \left[M\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n.$$

It will suffice to show that $\lim_{n \rightarrow \infty} n \log M(t/(\sigma\sqrt{n})) = t^2/2$; then we can take exponentials and use Theorem 3 to get the result. Call this limit L . Let $x = 1/\sqrt{n}$ in this limit; then we need to find

$$\lim_{x \rightarrow 0} \frac{\log M(tx/\sigma)}{x^2}.$$

This is of indeterminate form $0/0$, since $M(0) = 1$. Differentiating (l'Hopital's rule) gives

$$L = \lim_{x \rightarrow 0} \frac{\frac{M'(tx/\sigma) \frac{t}{\sigma}}{M(tx/\sigma) \frac{t}{\sigma}}}{2x}$$

and we can pull out a constant to get

$$L = \frac{t}{2\sigma} \lim_{x \rightarrow 0} \frac{M'(tx/\sigma)}{xM(tx/\sigma)}.$$

This is again indeterminate, of form 0 over 0 . Differentiating again gives

$$L = \frac{t}{2\sigma} \lim_{x \rightarrow 0} \frac{M''(tx/\sigma) \frac{t}{\sigma}}{M(tx/\sigma) + xM'(tx/\sigma) \frac{t}{\sigma}}.$$

Pulling out a constant and rearranging limits gives

$$L = \frac{t^2}{2\sigma^2} \frac{\lim_{x \rightarrow 0} M''(tx/\sigma)}{\lim_{x \rightarrow 0} M(tx/\sigma) + \frac{t}{\sigma} \lim_{x \rightarrow 0} xM'(tx/\sigma)} = \frac{t^2}{2\sigma^2} \frac{M''(0)}{M(0) + \frac{t}{\sigma} 0M'(0)}.$$

Now we recall $M^{(r)}(0) = E(X^r)$; thus $M(0) = E(1) = 1$, $M'(0) = E(X) = 0$, $M''(0) = E(X^2) = E(X)^2 + Var(x) = 0^2 + \sigma = \sigma$. This finally gives

$$L = \frac{t^2}{2\sigma^2} \frac{\sigma}{1 + 0} = \frac{t^2}{2}$$

which is what we wanted. □