

**PROBABILITY THEORY:
THE CENTRAL LIMIT THEOREM**

Note: *This work is intended for informative and educational purposes only.*

1. Introduction

The final issue of the year focuses on another important result in Probability Theory: the Central Limit Theorem. As with last month's issue, it is short and to the point—covering only the information required to get an understanding of the theorem.

2. Overview of the Theorem

Let's suppose we have a sequence of random samples, X_1, X_2, \dots, X_n drawn from a population of some kind. We will also state that the population has a mean, μ and variance $\sigma^2 < \infty$. Now suppose we take the mean \bar{X}_n of each of those samples. Then we say that the limiting distribution

$$\lim_{n \rightarrow \infty} \left(\frac{\bar{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \right) \quad (1)$$

is a standard normal distribution.

3. What does this mean?

Once again, this takes a little bit of unpacking to truly understand what we're saying. Let's consider an example. Suppose we have a standard uniform distribution,

$$X \sim \mathcal{U}(0, 10). \quad (2)$$

And suppose we take a random sample of, say, $N = 30$. That is, draw 30 random numbers from our uniform distribution X . Next we take the average of that sample, \bar{X} and record it. Great. Now let's repeat that. We take another randomly selected sample of $N = 30$, compute the mean and record it. Do it again. And again. Until we have a sufficiently large number of samples, for argument's sake I'll pretend we've done this 300 times.

If we were to produce a histogram of all of those sample means \bar{X}_n we would find they closely resemble a normal distribution, with mean approximately equal to the mean of our initial uniform distribution. That is, we would find an normal looking histogram centred around 5 (recall $\mathbb{E}(X) = \frac{1}{2}(a + b)$ for a Uniform distribution).

Now this example is slightly different to what we have described above as our overview of the theorem stated that the distribution tends towards a standard normal. Well that is true provided we properly normalise our sample means. Recall for to standardise a normally distributed random variable we must compute,

$$Z = \frac{X - \mu}{\sigma}. \quad (3)$$

For the central limit theorem, however, we need to include the correction term $\frac{1}{\sqrt{n}}$ in the denominator of (1). For our example though we didn't do this, so it only tends towards a normal distribution not a *standard* normal distribution.

4. Statement of the Theorem

Suppose we have a sequence of independent and identically distributed (i.i.d.) random variables $X_1, X_2, \dots, X_n, \dots$ with given expectation $\mathbb{E}(X_i) = \mu$, and given finite variance $\mathbb{V}ar(X_i) = \sigma^2 < \infty$. Then we have that as $n \rightarrow \infty$, the random variable defined as $\sqrt{n}(\bar{X}_n - \mu)$ converges to a normal distribution. That is,

$$\sqrt{n}(\bar{X}_n - \mu) \longrightarrow \mathcal{N}(0, \sigma^2). \quad (4)$$

An important point to note is that there are no conditions on which distribution the random variables must be from, merely that they all the same. This means the result holds for *any* distribution—which clearly shows how powerful a result this is.