

PROBABILITY THEORY

MATH 154

Unit 20: De Moivre-Laplace and Poisson

20.1. The central limit theorem is applied whenever we have to analyze data. If X_i are IID random variables in $\mathcal{L}^2$ with mean m and standard deviation σ , we can ask for the probability that S_n/n is within $\epsilon/\sqrt{n}$ to the mean m . The central limit helps us validate data by observing averages of experiments. Here is the application version:

Theorem 1. *The probability that S_n/n deviates more than $t\sigma/\sqrt{n}$ from the mean $m = E[X]$ can for large n be estimated by*

$$\frac{1}{\sqrt{2\pi}} \int_t^\infty e^{-x^2/2} dx .$$

Proof. Let $m = E[X]$ denote the mean of X_k and σ the standard deviation. Denote by X a random variable which has the standard normal distribution $N(0, 1)$. Write $X_n \sim Y_n$ if $X_n \rightarrow^d Y_n$ in distribution. By the central limit theorem

$$\frac{S_n - nm}{\sqrt{n}\sigma} \sim X .$$

Dividing both nominator and denominator by n gives $\frac{\sqrt{n}}{\sigma}(\frac{S_n}{n} - m) \sim X$ so that as distributions

$$\frac{S_n}{n} - m \sim X \frac{\sigma}{\sqrt{n}} .$$

But this means that we can estimate the deviation as $1 - F_{N(0,1)}(t)$, which is the expression we see in the theorem. $\square$

20.2. The term $\sigma/\sqrt{n}$ is called the **standard error**. The value t is the **z-score** and $1 - F_{N(0,1)}(t)$ the p-value. We will discuss a concrete case in class.

20.3. Let us look at the case of coin tossing. These are independent $\{0, 1\}$ -valued random variables with win probability $p \in (0, 1)$ was historically the starting point for the central limit theorem. The sum S_n has a **Binomial distribution** $B(n, p)$ of mean np and variance $np(1 - p)$. As we have just seen, the fact that

$$\lim_{n \rightarrow \infty} P\left[\frac{(S_n - np)}{\sqrt{np(1 - p)}} \leq x\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} dy$$

is a consequence of the central limit theorem. It has already been proven by de Moivre in 1730 in the case $p = 1/2$ and for general $p \in (0, 1)$ by Laplace in 1812.

Theorem 2 (DeMoivre-Laplace limit theorem). *If S_n have the Binomial distribution $B(n, p)$, then S_n^* converges in distribution to $N(0, 1)$.*

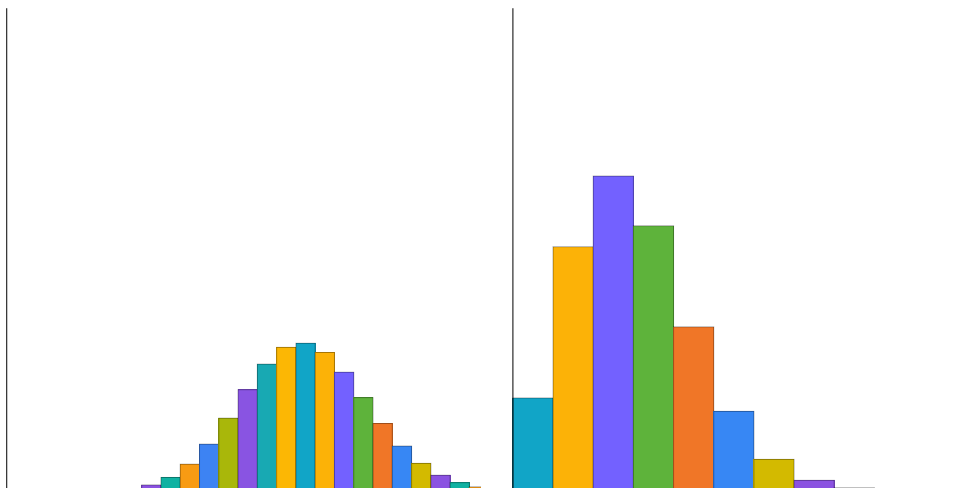


FIGURE 1. Binomial distribution $B(n, 0.3)$ to the left for $n = 50$. The Binomial distribution $B(n, 1/n)$ to the right for $n = 50$. The left will for $n \rightarrow \infty$ when rescaled will converge to the normal distribution. The right will converge to the Poisson distribution.

20.4. The **Poisson distribution** P_λ supported on $\mathbb{N} = \{0, 1, 2, 3, \dots\}$ is defined as

$$P[X = k] = e^{-\lambda} \frac{\lambda^k}{k!}.$$

Random variables with such distribution describe waiting times. Its moment generating function is

$$M_X(t) = \sum_{k=0}^{\infty} P[X = k] e^{tk} = e^{\lambda(1-e^{-t})}.$$

That the Poisson distribution is natural follows from:

Theorem 3 (Poisson limit theorem). *Let X_n be a $B(n, p_n)$ -distributed and suppose $np_n \rightarrow \lambda$. Then X_n converges in distribution to a random variable X with Poisson distribution with parameter λ .*

20.5. For the proof we need the last time used **compound interest statement** telling that if $a_n \rightarrow 0$, $b_n a_n \rightarrow c$ implies $(1 + a_n)^{b_n} \rightarrow e^c$. It is related to $(1 + c/n)^{1/n} \rightarrow e^c$.

Proof. We have to show that $P[X_n = k] \rightarrow P[X = k]$ for each fixed $k \in \mathbb{N}$.

$$\begin{aligned} P[X_n = k] &= \binom{n}{k} p_n^k (1 - p_n)^{n-k} \\ &= \frac{n(n-1)(n-2) \dots (n-k+1)}{k!} p_n^k (1 - p_n)^{n-k} \\ &\sim \frac{1}{k!} (np_n)^k \left(1 - \frac{np_n}{n}\right)^{n-k} \rightarrow \frac{\lambda^k}{k!} e^{-\lambda}. \end{aligned}$$

□