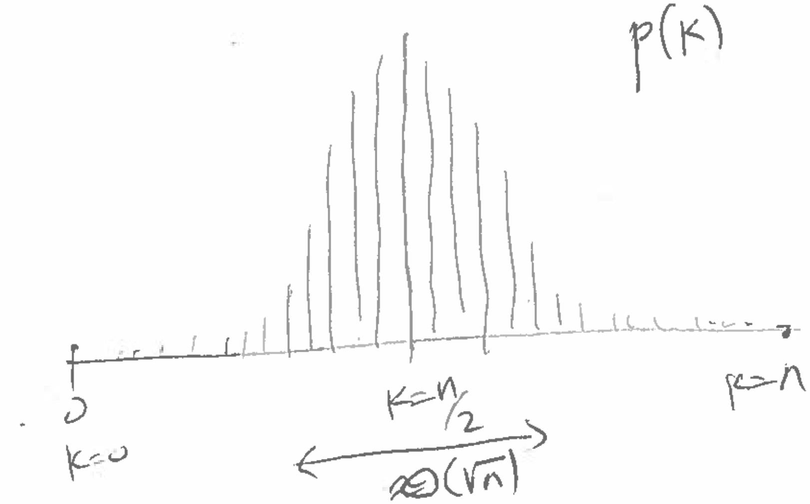
2023-02-1

Central Limit Theorem

\# of heads in n coin tosses is $\mathrm{Bin}(n, \tfrac{1}{2})$

$$p(k) = \binom{n}{k}\left(\tfrac{1}{2}\right)^k\left(\tfrac{1}{2}\right)^{n-k} = \binom{n}{k}\left(\tfrac{1}{2}\right)^n$$



Stirling: $n! \sim \sqrt{2\pi n}\,\left(\frac{n}{e}\right)^n$

$$p(k) = \frac{n!}{k!\,(n-k)!\,2^n}$$

$$p(k) \sim \frac{\sqrt{2\pi n}\,\left(\frac{n}{e}\right)^n}{\sqrt{2\pi k}\,\left(\frac{k}{e}\right)^k \sqrt{2\pi(n-k)}\,\left(\frac{n-k}{e}\right)^{n-k} 2^n}$$

$$p\left(\tfrac{n}{2}\right) \sim \frac{\sqrt{2\pi n}}{\left(\sqrt{2\pi n/2}\right)^2} = \frac{1}{\sqrt{\pi n}}$$

If $k \approx \frac{n}{2}$: $k = \frac{n}{2} + t$

$$p\left(\tfrac{n}{2}+t\right) \sim \frac{1}{\sqrt{\pi n}}\, e^{-t^2/2n}$$

valid if $n \gg 1$ $\quad t \ll n$

Similarly: $\mathrm{Poi}(\lambda)$: $f(\lambda+t) \approx \frac{1}{\sqrt{2\pi\lambda}} e^{-t^2/2\lambda}$

Gamma R.V. also has normal approx.

<u>Central Limit Theorem</u> (CLT):

Let $X_1, X_2 \ldots$ be iid R.V.s with $\mu = EX_i$ and $\sigma^2 = \mathrm{Var}(X_i)$. Let $S_n = \sum_{i=1}^{n} X_i$. Then

$$\underbrace{\frac{S_n - \mu n}{\sqrt{n}\,\sigma}}_{Z_n} \xrightarrow{D} N(0,1).$$

$$\left[P(a \le Z_n \le b) \xrightarrow[n\to\infty]{} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \qquad \forall a,b \right]$$

$Z_n \approx N(0,1)$

$S_n = \mu n + \sigma\sqrt{n}\, Z_n$ LLN: $\frac{1}{n} S_n \longrightarrow \mu$

<u>proof of CLT</u>: by continuity thm, need to show that

$$\phi_{Z_n}(t) \xrightarrow[n\to\infty]{} e^{-t^2/2}.$$

Let $Y_n = X_n - \mu$ so $EY_n = 0$ $Var(Y_n) = \sigma^2$

$$Z_n = \frac{Y_1 + Y_2 + \cdots Y_n}{\sigma\sqrt{n}} \qquad \phi_{Z_n}(t) = \phi_{Y_1+Y_2 \cdots + Y_n}\left(\frac{t}{\sigma\sqrt{n}}\right)$$

$$= \left[\phi_Y\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n$$

$$\phi_Y(s) = 1 + isEY + \frac{i^2 s^2}{2!} EY^2 + O(s^3)$$

$$\phi_Y(s) = 1 + \frac{s^2}{2} \cdot \sigma^2 + err$$

$$\left[\phi_Y\left(\frac{t}{\sigma\sqrt{n}}\right)\right]^n = \left[1 - \frac{t^2\sigma^2}{\sigma^2 \cdot n \cdot 2} + err\right]^n \longrightarrow e^{-t^2/2} \quad \text{as needed.}$$

$\square$

---

note

⊛ If $EX = \infty$ or not defined then $\frac{S_n}{n}$ has no limit

⊛ If $EX = \mu$ but $Var(X) = \infty$ then $\frac{S_n}{n} \longrightarrow \mu$

but $|S_n - n\mu| \gg \sqrt{n}$

E.g. want 50 new students.
Historically 40% of admitted students accept the offer.
If 100 students admitted, what is P(more than 50 come)?

$X \sim Bin(100, 0.4)$ need $P(X > 50)$

X is a sum of 100 $Bern(0.4)$

$E\, Bern(p) = p = 0.4$

$Var(Bern(p)) = p(1-p) = 0.2$

CLT $\dfrac{X - 100p}{\sqrt{100\, p(1-p)}} \approx N(0,1)$

$$Z = \frac{X - 40}{\sqrt{24}} \approx N(0,1)$$

$$X > 50 \iff \frac{X-40}{\sqrt{24}} > \frac{50-40}{\sqrt{24}} = \frac{10}{\sqrt{24}} = 2^{+} \text{ a bit} = 2.04$$

$$P(Z > 2^{+}) = 1 - \phi(2.04)$$

$$\phi(t) = P\big(N(0,1) \le t\big)$$