

Math 340: Lec 16 Big Ideas Journal (Random walks continued; central limit theorem)

Asa Royal (ajr74)

March 5, 2024

Random walks

Theorem 1 (Ballot theorem). Consider $N_n^+(0, b)$: paths from $S_0 = 0$ to $S_n = b$ for which $S_k > 0 \forall k \in \{1, \dots, n\}$. We can think of S_k as how many votes ahead candidate X is over candidate Y on election night.

$$\forall b \neq 0, N_n^+(0, b) = \frac{|b|}{n} N_n(0, b)$$

$|b|/n$ is the fraction of paths that don't touch the x -axis (never go negative).

Central limit theorem

CLT applied to random walks

Example (CLT random walks). Consider the following probability measure (bakes in equal parity)

$$\mathbb{P}_n^0(S_{2n} = 2k) = \binom{2n}{n+k} 2^{-(2n)}$$

Let X_k be a random variable denoting how much we move on step k .

$$\mathbb{E}[S_n] = \mathbb{E}\left[\alpha + \sum_{k=1}^n X_k(\omega)\right] = \sum_{k=1}^n \mathbb{E}[X_k] = 0$$

$$\text{Var}(S_n) = \sum_{k=1}^n \text{Var}(X_k) = \sum_{k=1}^n \mathbb{E}(|X_k - \mu|^2) = \sum_{k=1}^n \mathbb{E}(|X_k - 0|^2) = \sum_{k=1}^n 1 = n$$

$$SD(S_n) = \sqrt{n}$$

Remark. The CLT suggests us that we shouldn't be surprised if $S_n \approx O(\sqrt{n})$.

Theorem 2 (Central Limit Theorem for random walks). For any α, β

$$\lim_{n \rightarrow \infty} \mathbb{P}_n^\alpha \left(\alpha \leq \frac{S_n}{\sqrt{n}} \leq \beta \right) = \int_\alpha^\beta \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy$$

Remark. The CLT lets us bound how far away we expect a random walker to wander from the mean of the walk.

General CLT

Theorem 3 (Central Limit Theorem). Suppose X_1, \dots, X_n is a sequence of i.i.d random variables with $\mu = \mathbb{E}[X_i]$, $\sigma^2 = \text{Var}(X_i)$ and $\mathbb{E}[X_i^4] < \infty$. For any α, β

$$\lim_{n \rightarrow \infty} \mathbb{P}_n^\alpha \left(\alpha \leq \frac{(X_1 + \dots + X_n) - \mu_n}{\sqrt{n\sigma^2}} \leq \beta \right) = \int_\alpha^\beta \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy = \Phi(\beta) - \Phi(\alpha)$$

Remark. Note that the CLT for random walks is a particular case of the general CLT where S_n is a sum of random step variables with mean zero and variance 1.

Example (Using CLT to bound p-coin head count). Let Z_n be the nubmer of heads we see in n tosses. The marignal distribution of Z_n is given by the binomial distribution. As we know, $Z_n = X_1 + \dots + X_n$ where X_j is a bernoulli random variable. $\mathbb{E}[X_j] = p$ and $\text{Var}(X_j) = p(1 - p)$. Per the CLT:

$$\mathbb{P} \left(\alpha \leq \frac{(X_1 + \dots + X_n) - np}{\sqrt{np(1 - p)}} \leq b \right) = \int_{\alpha}^{\beta} \frac{e^{-\frac{y^2}{2}}}{\sqrt{2\pi}} dy$$

Also, note that the CLT expression gives us

$$\begin{aligned} & \mathbb{P} \left(\alpha \sqrt{np(1 - p)} \leq (X_1 + \dots + X_n) - np \leq \beta \sqrt{np(1 - p)} \right) \\ &= \mathbb{P}(Z_n \in (np + \alpha \sqrt{np(1 - p)}, np + \beta \sqrt{np(1 - p)}) \end{aligned}$$

i.e. the probabiliy that Z is within α and β standard deviations of its mean.