(c) **Prop** (Stability of normal distribution)

If $X_i \sim N(\mu_i, \sigma_i^2)$ are independent then $\sum_{i=1}^n X_i \sim N\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right)$

Proof ① ($\mu_i = 0$, $\sum \sigma_i^2 = 1$): Assume $Z_i \sim N(0,1)$ are independent and $\sum_1^n \sigma_i^2 = 1$.

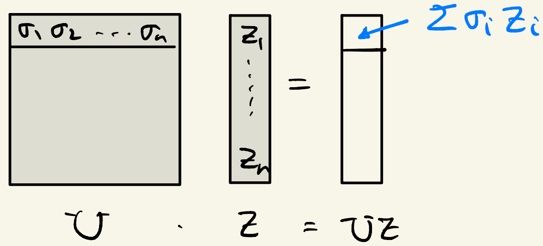
Then $\sum_{i=1}^n \sigma_i Z_i \sim N(0,1)$

$Z := (Z_1, \dots, Z_n) \sim N(0, I_n)$.

Rotation invariance $\Rightarrow$ $\boxed{UZ \sim N(0, I_n)}$ $\forall U \in O(n)$.

Find a matrix $U \in O(n)$ with first row $(\sigma_1, \dots, \sigma_n)$ $\exists$ by (Gram–Schmidt)

$\Rightarrow$



$\Rightarrow \sum \sigma_i Z_i \sim N(0,1)$.

② (General case): by dilation & translation (DIY). $\square$

HW: Prove by convolution

Cor The coordinates of a r.v. $X \sim N(\mu, \Sigma)$ are normal

By def, $X = AZ + \mu$ for some $Z \sim N(0, I_n)$, $A \in \mathbb{R}^{n \times n}$ invertible, $\mu \in \mathbb{R}^n$

$\Rightarrow X_i = \sum_{j=1}^n A_{ij} Z_j + \mu_i \sim N\left(\mu_i, \sum_{j=1}^n A_{ij}^2\right)$

Remark There are other stable distributions. Example: Cauchy distribution

$X \sim$ Cauchy has pdf $f(x) = \frac{1}{\pi(1+x^2)}$, $x \in \mathbb{R}$.

$\mathbb{E}X$ does NOT exist (heavy tail)

THM (DIY): If $X_1, \dots, X_n \sim$ Cauchy iid then $\frac{1}{n}\sum_1^n X_i \sim$ Cauchy

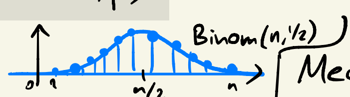
All stable distributions other than normal have ∞ variance (By CLT) $\Rightarrow$ Heavy-tailed

(d) Binomial distribution Let $X_i \sim Ber(p)$ be iid.

$S_n := X_1 + \dots + X_n \sim$ Binom(n,p) = # successes in n independent trials

Probability mass function:

$P\{S_n = k\} = \binom{n}{k} p^k (1-p)^{n-k}$,



$k = 0, 1, \dots, n$.

Mean: $\mathbb{E}S_n = n \cdot \mathbb{E}X_1 = np$

Variance: $Var(S_n) = n \cdot Var(X_1) = np(1-p)$

THE WEAK LAW OF LARGE NUMBERS (WLLN)

Def $X_n \to X$ in probability if $\forall \varepsilon > 0$:

$$\mathbb{P}\{|X_n - X| > \varepsilon\} \to 0 \text{ as } n \to \infty$$

THM (Weak LLN) Let $X_1, X_2, \dots$ be independent r.v's with $\mathbb{E}X_i = \mu$ and $\sup_n \operatorname{Var}(X_i) < \infty$. Then $S_n := X_1 + \dots + X_n$ satisfies

$$\frac{S_n}{n} \to \mu \text{ in probability}$$

Proof $\operatorname{Var}\left(\frac{S_n}{n}\right) = \frac{1}{n^2}\operatorname{Var}(S_n) = \frac{1}{n^2}\left[\underbrace{\operatorname{Var}(X_1)}_{\leq C} + \dots + \underbrace{\operatorname{Var}(X_n)}_{\leq C}\right] \leq \frac{C}{n}.$ where $C = \sup_n \operatorname{Var}(X_i)$

Chebyshev's inequality: $\forall \varepsilon > 0$

$$\mathbb{P}\left\{\left|\frac{S_n}{n} - \mu\right| > \varepsilon\right\} \leq \frac{\operatorname{Var}(S_n/n)}{\varepsilon^2} = \frac{C}{n\varepsilon^2} \to 0 \text{ as } n \to \infty. \qquad \square$$

Remarks
1. It's enough if X_i are uncorrelated
2. We get a quantitative rate of error's decay

   Usually, it is better: $e^{-c(\varepsilon)n}$ ("concentration inequalities")

Application:

continuous functions $[a,b] \to \mathbb{R}$

Weierstrass Approximation Thm: Polynomials are dense in $C[a,b]$. That is:

$\forall f \in C[a,b]$ $\exists$ sequence of polynomials f_n such that

$$\lim_n \underbrace{\|f_n - f\|_\infty}_{\sup_{x\in[a,b]} |f_n(x) - f(x)|} = 0.$$

Probabilistic proof WLOG $[a,b] = [0,1]$.

- Will show that Bernstein polynomials work:

$$f_n(x) := \sum_{k=0}^{n} \underbrace{\binom{n}{k} x^k (1-x)^{n-k}}_{=\,\mathbb{P}\{S_n = k\}} f\left(\frac{k}{n}\right).$$

- Fix $x \in [0,1]$, let $S_n \sim \text{Binom}(n,x)$ $\Rightarrow$

$$\mathbb{E} f\left(\frac{S_n}{n}\right) = \sum_{k=0}^{\infty} \mathbb{P}\{S_n = k\} f\left(\frac{k}{n}\right) = f_n(x).$$

$$\Rightarrow |f_n(x) - f(x)| = \left|\mathbb{E} f\left(\frac{S_n}{n}\right) - f(x)\right| \overset{\text{Jensen}}{\leq} \mathbb{E}\left|f\left(\frac{S_n}{n}\right) - f(x)\right| \ominus$$

A continuous function $f: [0,1] \to \mathbb{R}$ is:

(a) uniformly continuous $\Rightarrow$ $\forall \varepsilon > 0$ $\exists \delta > 0$: $|x-y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon$

(b) bounded $\Rightarrow$ $\|f\|_\infty \leq M < \infty$.

$$\ominus\ \mathbb{E}\underbrace{\left|f\left(\frac{S_n}{n}\right) - f(x)\right|}_{\wedge\ \varepsilon} \mathbb{1}_{\{|\frac{S_n}{n} - x| < \delta\}} + \mathbb{E}\underbrace{\left|f\left(\frac{S_n}{n}\right) - f(x)\right|}_{\wedge\ 2M} \mathbb{1}_{\{|\frac{S_n}{n} - x| \geq \delta\}}$$

$$\leq \varepsilon + 2M \cdot \underbrace{\mathbb{P}\left\{\left|\frac{S_n}{n} - x\right| \geq \delta\right\}}_{\wedge \text{ Chebyshev}}$$

$$\frac{\operatorname{Var}(S_n/n)}{\delta^2} = \frac{\operatorname{Var}(S_n)}{n^2\delta^2} = \frac{nx(1-x)}{n^2\delta^2} \leq \frac{1}{n\delta^2}$$

Binom(n,x)

$$\leq \varepsilon + \frac{2M}{n\delta^2}.$$

- Hence, we proved: $\forall \varepsilon > 0$ $\exists \delta > 0$: $\|f_n - f\|_\infty \leq \varepsilon + \frac{2M}{n\delta^2}$

Let $n \to \infty$ $\Rightarrow$ $\forall \varepsilon > 0$: $\limsup_n \|f_n - f\|_\infty \leq \varepsilon$

$$\Rightarrow \limsup_n \|f_n - f\|_\infty = 0. \quad \square$$

HW Quantitative version:

$$\|f_n - f\|_\infty \leq \frac{1}{\sqrt{n}} \|f\|_{Lip}$$