Section 15

Wednesday, December 7, 2022 8:48 PM

Correction from last time - we can use Chebyshev for bounds.

(Ex) Suppose the mean on a probability exam is 75.

With just this information,

(i) Find an upperbound for probability that someone scored above 90.

(iii) Suppose you gain some information - specifically, the standard deviation was 5. Answer (1)

Before stating CLT, a warmup:

(Q) Suppose $X_1, X_2 \dots X_n$ are iid w/ $\mathbb{E}X_i = \mu$

$Var(X_i) = \sigma^2$.

Consider $S_n = X_1 + \dots + X_n$, "sample total"

$\bar{X} = \frac{S_n}{n}$, "sample mean"

(a) Express $\mathbb{E}S_n$ and $\mathbb{E}\bar{X}$ in terms of above defined parameters.

(b) Similarly, compute $Var(S_n)$ and $Var(\bar{X})$.

(Df) (Convergence in distribution)

Let $X_1, X_2 \dots$ be random variables. If there's random variable X s.t., for all $x \in \mathbb{R}$, $\lim_{n\to\infty} \mathbb{P}[X_n \le x] = \mathbb{P}[X \le x]$, then X_n converges to X in distribution, or $X_n \to_{\mathcal{D}} X$.

(We actually require this for only $x \in \mathbb{R}$ such that $\mathbb{P}[X \le x]$ is continuous at x)

(Rm) So, if we have $X_n \to_{\mathcal{D}} X$, we may assert that $\mathbb{P}[X \le x] \approx \mathbb{P}[X_n \le x]$, i.e. approximate X_n's probability w/ X's.

Central Limit Theorem

(Th) Say, $X_1, X_2 \dots$ are iid w/ $\mathbb{E}X_i = \mu < \infty$, $Var(X_i) = \sigma^2 < \infty$.

Then, consider $S_n = \Sigma X_i$. We have:

$$\frac{S_n - n\mu}{\sqrt{n}\sigma} \to_{\mathcal{D}} N(0,1)$$

$\Updownarrow$ (Equivalently)

$$\frac{\frac{S_n}{n} - \mu}{\sigma/\sqrt{n}} \to_{\mathcal{D}} N(0,1)$$

In words, $\mathbb{P}\left[\frac{\frac{S_n}{n} - \mu}{\sigma/\sqrt{n}} < z \right] \approx \phi(z)$.

(A canonical example)

$\boxed{\text{Ques}}$ Say, 20 of us toss a coin 10 times, and count the number of heads we get. Let X_i = # of heads that i^{th} person gets. Let $\bar{X} = \sum_{i=1}^{20} X_i/20$, "sample mean".

(i) Compute $\mathbb{E}\bar{X}$ and $\mathrm{Var}(\bar{X})$.

(ii) Approximate, using CLT, $\mathbb{P}[\bar{X} \le 10]$.

(Playing roulette)

$\boxed{\text{Ques}}$.
- A roulette wheel has 18 red and 18 black slots, and 2 green slots.
- Players can bet \$1 that ball lands in red (or black) slot, and win \$1 if it does.

(a) Let X_i = winnings on the i^{th} play. What is pmf of X_i?

(b) Compute $\mathbb{E}X_i$, $\mathrm{Var}(X_i)$ (You may approximate $\mathrm{Var}(X_i)$ to 1 significant digit)

(c) Let $S_n = X_1 + \dots + X_n$ count winnings up till your n^{th} attempt. If $n = 19$, compute $\mathbb{P}[S_n \geq 0]$ i.e. you don't have net loss.

(Crazy big numbers)

$\boxed{\text{Ques}}$ Say, $U_1 \dots U_{2000} \sim \mathrm{unif}(0,1)$ iid. Approximate $\mathbb{P}[1980 \le S_{2000} \le 2020]$, where $S_{2000} = \sum_{i=1}^{2000} U_i$.

Jensen's inequality

Jensen's inequality

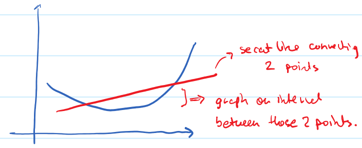
(Df) (Convexity)

A function $g: I \to \mathbb{R}$, where I is some interval in $\mathbb{R}$, is <u>convex</u> if

for all $0 \le \lambda \le 1$ and $x_1, x_2 \in I$,

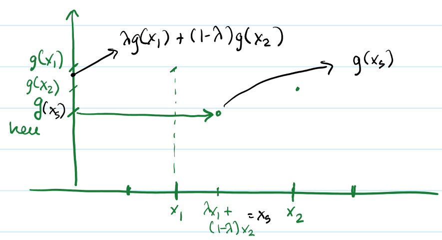
$$g(\lambda x_1 + (1-\lambda)x_2) \le \lambda g(x_1) + (1-\lambda) g(x_2).$$

"secant line connecting 2 points lies <u>above</u> the graph on the interval between those 2 points"



Choose $\lambda > 0$,

$x_1, x_2 \in I$.



(Th) If X is an r.v. with φ convex on the support of X, $\varphi(\mathbb{E}[X]) \le \mathbb{E}[\varphi(X)]$. Write (pf) from Gabe's book + 1 application.

(Pf) (A1) Say, X is finite discrete r.v., w/ support $\{x_1 \dots x_n\}$

(A2) By induction, <u>we first show:</u>

(Lm) if $g: I \to \mathbb{R}$ is convex, then for $\lambda_1 \dots \lambda_m$ s.t. $\sum \lambda_i = 1$,

and $x_1, \dots x_m \in I$, $g(\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_m x_m) \le \lambda_1 g(x_1) + \dots + \lambda_m g(x_m)$

(Pf) By induction on m:

(1) <u>$m = 2$</u>

Then, letting $\lambda := \lambda_1$, we see $\lambda_2 + \lambda_1 = 1 \Rightarrow \lambda_2 = 1 - \lambda$, and the claim $g(\lambda_1 x_1 + \lambda_2 x_2) \le \lambda_1 g(x_1) + \lambda_2 g(x_2)$ follows since g is convex.

(2) <u>Assume true for $m = k \in \mathbb{N}$</u>

(3) Then, for $m = k+1$ ...

(a) Say, $0 < \lambda_1 \dots \lambda_{k+1} < 1$ s.t. $\sum_{i=1}^{k+1} \lambda_i = 1$

(b) $x_1 \dots x_{k+1} \in I$.

(c) <u>To show:</u> $g(\lambda_1 x_1 + \dots + \lambda_k x_k + \lambda_{k+1} x_{k+1}) \le \lambda_1 g(x_1) + \dots + \lambda_k g(x_k) + \lambda_{k+1} g(x_{k+1})$

We know $\lambda_1 + \dots + \lambda_k = 1 - \lambda_{k+1} > 0$. Then, we rewrite:

$$\underline{\lambda_1 x_1 + \dots + \lambda_k x_k} + \lambda_{k+1} x_{k+1} = (1 - \lambda_{k+1}) \underbrace{\boxed{\frac{\lambda_1 x_1 + \dots + \lambda_k x_k}{1 - \lambda_{k+1}}}}_{= y \text{, say.}} + \lambda_{k+1} x_{k+1}$$

I claim that $y \in I$, (because I is convex, and y is convex combo. of points in I)

Then, since g is convex, we know $g((1-\lambda_{k+1}) y + \lambda_{k+1} x_{k+1}) \le (1-\lambda_{k+1}) g(y) + \lambda_{k+1} g(x_{k+1})$

$$= (\lambda_1 + \dots + \lambda_k)\, g\left(\frac{\lambda_1 x_1 + \dots + \lambda_k x_k}{\lambda_1 + \dots + \lambda_k}\right) + \lambda_{k+1} g(x_{k+1})$$

$$\le \lambda_1 g(x_1) + \dots + \lambda_k g(x_k) + \lambda_{k+1} g(x_{k+1}), \text{ as desired.}$$

(A3) Now, note that $\mathbb{E}X = p_1 x_1 + \dots + p_n x_n$, where $p_i = \mathbb{P}[X = i]$.

We know $\sum_{i=1}^{n} p_i = 1$, $0 < p_i < 1$, and so we may apply above lemma:

$\varphi(\mathbb{E}X) = \varphi(p_1 x_1 + \cdots + p_n x_n) \leq p_1 \varphi(x_1) + \cdots + p_n \varphi(x_n) = \mathbb{E}[\varphi(X)]$, as desired.