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.

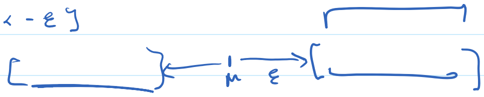
With just this information,

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

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

Chebyshev $\mathbb{P}[\underbrace{|X-\mu|>\varepsilon}_{\overset{\shortparallel}{[X-\mu>\varepsilon]\cup[X-\mu<-\varepsilon]}}] \le \frac{Var(X)}{\varepsilon^2}$, where $\varepsilon > 0$, $\mathbb{E}X = \mu$

$\mathbb{P}[\underbrace{X-\mu>\varepsilon}_{\uparrow}] \le$



Before stating CLT, a warmup:

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

$Var(X_i) = \sigma^2$. ←

Consider $S_n = X_1 + \ldots + X_n$, "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.

- $\mathbb{E}S_n = \mathbb{E}\sum_{i=1}^{n} X_n \overset{\downarrow}{=} n \cdot \mathbb{E}X_i = n\mu$ ←
- $\mathbb{E}\bar{X} = \mathbb{E}\frac{S_n}{n} = \boxed{\mu}$ ⟸

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

- $Var(\sum_{i=1}^{n} X_i) = \sum_{i=1}^{n} Var(X_i) = n\sigma^2$ ←
- $Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) = \frac{1}{n^2}\sum_{i=1}^{n} Var(X_i) = \frac{n\sigma^2}{n^2} = \boxed{\frac{\sigma^2}{n}}$ ←

(Df) (Convergence in distribution)

Let $\overset{F_1}{X_1}, \overset{F_2}{X_2} \ldots$ be random variables. If there's random variable X s.t., $\xrightarrow{} F := \mathbb{P}[X \le x]$

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

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

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

⇕ (Equivalently)

"normalized sample mean" ← $\frac{\frac{S_n}{n} - \mu \ \times n}{\sigma/\sqrt{n} \ \times n} \to_{\mathcal{D}} N(0,1)$

distribution of standard

Sample mean" $\leftarrow \dfrac{\frac{\ }{n} - \mu}{\sigma/\sqrt{n}} \times n \longrightarrow_{D} \underbrace{N(0,1)}$

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

$\uparrow$ distribution of standard normal

(A canonical example)

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

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

$X_i \sim \text{Binomial}(10, 1/2)$, $\mathbb{E}X_i = \underbrace{5}$ $Var(X_i) = 10 \cdot 1/2^2 = \frac{10}{4} = \overline{2.5}$

- $\underline{\mathbb{E}\bar{X} = 5}$
- $Var(\bar{X}) = Var\left(\frac{\sum_{i=1}^{20} X_i}{20}\right) = \frac{1}{20^2} \cdot 20 \cdot Var(X_i) = \frac{1}{20} \cdot \frac{10}{4} = \underline{\frac{1}{8} = Var(\bar{X})} \leftarrow$

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

$$\mathbb{P}\left[\frac{\bar{X} - \mu}{\sigma/\sqrt{n}} < \frac{10 - \mu}{\underbrace{\sigma/\sqrt{n}}_{\sqrt{Var(X_i)}}}\right] \approx \phi\left(\frac{10-5}{\sqrt{1/8}}\right) = \phi\left(\frac{5}{\sqrt{\frac{1}{8}}}\right) \approx 1$$

(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?

$X_i \in \{-1, 1\}$. $\mathbb{P}[X_i = 1] = \frac{18}{38}$, $\mathbb{P}[X_i = -1] = \frac{20}{38}$.

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

- $\underline{\mathbb{E}X_i} = \sum_{x_i \in \{-1,1\}} x_i \, \mathbb{P}[X_i = x_i] = \frac{18}{38} - \frac{20}{38} = \frac{-2}{38} = \frac{-1}{19}$.
- $\underline{Var(X_i)} = \sum_{x_i \in \{-1,1\}} (x_i - \mathbb{E}X_i)^2 \, \mathbb{P}[X_i = x_i] \approx \underline{1}$ $(= 0.9972..)$

(c) Let $\underline{S_n} = X_1 + \ldots + X_n$ count winnings up till your $\underline{n^{th}}$ attempt.

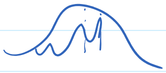
If $\underline{n = 19}$, compute $P[S_n \ge 0]$ i.e. you don't have net loss.

$$\mathbb{P}\left[\frac{S_n - \mathbb{E}S_n}{(Var(S_n))^{1/2}} \ge \frac{0+1}{\sqrt{19}}\right]$$

Now, $\mathbb{E}S_n = 19 \times \left(\frac{-1}{19}\right) = -1$

$Var(S_n) = n\,Var(X_i) = 19$

$$\approx \mathbb{P}\left[Z \ge \frac{1}{\sqrt{19}}\right] = 1 - \underbrace{\phi\left(\frac{1}{\sqrt{19}}\right)}_{>0.5} < 0.5$$



(Crazy big numbers)

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

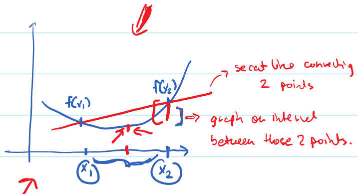
Jensen's inequality

(Df) (Convexity)

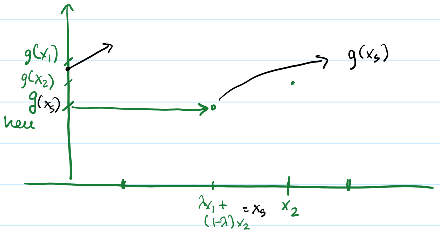
A function $g: I \to \mathbb{R}$, where I is some interval in $\mathbb{R}$, is **convex** if for all $0 \le \lambda \le 1$ and $x_1, x_2 \in I$,

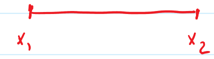
$$g(\lambda x_1 + (1-\lambda)x_2) \le \lambda g(x_1) + (1-\lambda) g(x_2).$$

"secant line connecting 2 points lies above the graph on the interval between those 2 points"



Choose $\lambda > 0$, $x_1, x_2 \in I$.





$= \lambda x_1 + (1-\lambda)x_2$

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

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

(A2) By induction, we first show:

(Lm) if $g: I \to \mathbb{R}$ is convex, then for $\overset{\curvearrowright p_i}{\overbrace{\lambda_1 \dots \lambda_m}}$ s.t. $\overbrace{\sum \lambda_i = 1}$, and $\overbrace{x_1, \dots x_m} \in I$, $g(\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_m x_m) \le \lambda_1 g(x_1) + \dots + \lambda_m g(x_m)$

(Pf) By induction on m:

(1) $m = 2$

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

(2) Assume true for $m = k \in \mathbb{N}$ $\Leftarrow$

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

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

$$\underline{\lambda_1 x_1 + \dots + \lambda_k x_k} + \lambda_{k+1} x_{k+1} = (1 - \lambda_{k+1}) \underbrace{\boxed{\frac{[\lambda_1 x_1 + \dots + \lambda_k x_k]}{1 - \lambda_{k+1}}}}_{} + \lambda_{k+1} x_{k+1}$$

$= y$, say.

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

$$g((1-\lambda_{k+1}) y + \lambda_{k+1} x_{k+1}) \le (1-\lambda_{k+1}) g(y) + \lambda_{k+1} g(x_{k+1})$$

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

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

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

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

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

$\boxed{\underline{\varphi(\mathbb{E}X)}} = \underline{\varphi(p_1x_1 + \cdots + p_nx_n)} \underline{\leq} \underline{p_1\varphi(x_1) + \ldots + p_n\varphi(x_n)} = \boxed{\mathbb{E}[\underline{\varphi(X)}]}$, as desired.