STAT 417. Lecture 3.

American's ~~expected~~ life expectation is 79 (years). ~~life~~

What is p(you can live longer than 100) ?

~~Today we will use Chebyshev's inequality to solve this problem.~~

Suppose X is a non-negative random variable. $E(X) = \mu$.

<u>thm 1</u> (markov inequality). For any real number $a > 0$.

$$P(X \geq a) \leq \frac{E(X)}{a}$$

<u>proof</u>: $E(X) = \int_0^\infty x f(x)\,dx \geq \int_a^\infty x f(x)\,dx \geq a \int_a^\infty f(x)\,dx$

$$= a\, p(X \geq a).$$

so $p(X \geq a) \leq \frac{E(X)}{a}$.

Back to our motivating example. let X = life time of a human being.

$$p(X \geq 100) \leq \frac{E(X)}{100} = 0.79.$$

<u>thm 2</u> (Chebyshev inequality). let $\mu = E(X)$. then. for any $a > 0$.

$$P(|X - \mu| \geq a) \leq \frac{Var(X)}{a^2}.$$

proof: $p(|X-\mu| \geq a) = p(|X-\mu|^2 \geq a^2) \leq \frac{E((X-\mu)^2)}{a^2} = \frac{Var(X)}{a^2}.$

If we know var(~~life~~ human's life time) $= 10$.

then. $p(X \geq 100) = p(X - 79 \geq 21) \leq p(|X - 79| \geq 21) \leq \frac{Var(X)}{21^2}$

$$= 0.023.$$

E.g. 1. If $X \sim$ Exponential(3). Then $E(X) = \frac{1}{3}$. $Var(X) = \frac{1}{9}$.

What is $\cancel{P(X \geq \frac{5}{6})}$ an upper bound for $P(X \geq \frac{5}{6})$?

Sol: using markov inequality,

$$P(X \geq \tfrac{5}{6}) \leq \frac{E(X)}{5/6} = \frac{1/3}{5/6} = \frac{2}{5}.$$

using chebychev inequality.

$$P(X \geq \tfrac{5}{6}) = P(|X - \tfrac{1}{3}| \geq \tfrac{1}{2}) \leq \frac{Var(X)}{(\frac{1}{2})^2} = \frac{4}{9}.$$

$\frac{2}{5} < \frac{4}{9}$, so. $P(X \geq \frac{5}{6}) \leq \frac{2}{5}$ is sharper.

E.g. 2. If $X \sim N(0,1)$. Then $E(X) = 0$, $Var(X) = 1$. $E(|X|) \approx 0.80$.

what is an upper bound for $P(|X| \geq 5)$?

chebychev: $P(|X| \geq 5) \leq \frac{Var(X)}{5^2} = \frac{1}{25} = 0.04$ (better).

markov: $P(|X| \leq 5) \leq \frac{E(|X|)}{5} = 0.16$.

Ex. 1. If $X \sim$ Binomial$(100, \frac{1}{2})$, so $E(X) = 50$, $Var(X) = 25$

$E(|X-50|) \approx 3.93$

Find an upper bound for $P(|X - 50| \geq 10)$.

sol. chebyshev: $P(|X - 50| \geq 10) \leq \frac{25}{100} = \frac{1}{4} = 0.25$. (better).

Markov: $P(|X-50| \geq 10) \leq \frac{E(|X-50|)}{10} = \frac{3.93}{10} = 0.393$.

~~Suppose~~ sample mean: Suppose $X_1, X_2, \ldots, X_n$ are i.i.d random variables. let $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$. then $\bar{X}$ is called sample mean

Question: ① what is the limit of $\bar{X}$?

② what is the distribution of $\bar{X}$?

To answer Q①, we have the following law of large numbers.

~~Def 3~~. convergen in probability: a random variable Y_n is said to converge in probability to Y, denoted $Y_n \xrightarrow{p} Y$, if for any $\varepsilon > 0$, $\lim_{n \to \infty} P(|Y_n - Y| \geq \varepsilon) = 0$.

E.g. ~~4~~. If. $Y_n = 1 - \frac{1}{n}$, $Y = 1$, show that $Y_n \xrightarrow{p} Y$.

For any $\varepsilon > 0$.

$P(|Y_n - Y| \geq \varepsilon) = P(\frac{1}{n} \geq \varepsilon) = 0$, if $n \to \infty$.

so. $Y_n \xrightarrow{p} Y$.

Thm3. (law of large number). Let $X_1, X_2, \ldots, X_n$ be i.i.d random variables. $E(X_i) = \mu$, $Var(X_i) = \sigma^2 < \infty$. Then $\bar{X}_n \xrightarrow{p} \mu$.

proof: For any $\varepsilon > 0$.

$$P(|\bar{X} - \mu| \geq \varepsilon) \leq \frac{Var(\bar{X})}{\varepsilon^2}.$$

while $Var(\bar{X}) = Var\left(\frac{X_1 + X_2 + \cdots + X_n}{n}\right) = \frac{1}{n^2}\left(Var(X_1) + \cdots + Var(X_n)\right) = \frac{1}{n^2} \cdot n\sigma^2$

$= \frac{1}{n}\sigma^2$.

So. $P(|\bar{X} - \mu| \geq \varepsilon) \leq \frac{\sigma^2}{n\varepsilon^2} \to 0$. as $n \to \infty$.

So. $\bar{X} \xrightarrow{p} \mu$.

E.g. 3. If $Y_1 = Y_2 = Y_3 = \cdots = Y$. then. $Y_n \xrightarrow{P} Y$.

proof: For any $\varepsilon > 0$.

$P(|Y_n - Y| \geq \varepsilon) = P(0 \geq \varepsilon) = 0$.

E.g. 5. If $Y_n \sim$ Exponential(n), $Y = 0$ then $Y_n \xrightarrow{P} Y$.

proof. For any $\varepsilon > 0$.

$P(|Y_n - 0| \geq \varepsilon) = P(Y_n \geq \varepsilon) = \int_{\varepsilon}^{\infty} n e^{-nx} \, dx = e^{-n\varepsilon} \to 0$.

E.X. 2. If ~~Y_n U~~ $U \sim$ unif$(0,1)$. $Y_n = U^n$.

then $Y_n \xrightarrow{P} 0$.

proof: For any $\varepsilon > 0$.

$P(|Y_n - 0| \geq \varepsilon) = P(Y_n \geq \varepsilon) = P(U \geq \varepsilon^{\frac{1}{n}}) = 1 - \varepsilon^{\frac{1}{n}} \to 0$.

Two answer Q 2:

Convergen in distribution: $Y_n \xrightarrow{d} Y$. Y_n converges to Y in distribution, denoted as $Y_n \xrightarrow{d} Y$, if the cdf of Y_n converges to the cdf of Y.

namely, for any y. such that $P(Y = y) = 0$,

$$\lim_{n \to \infty} P(Y_n \le y) = P(Y \le y).$$

thm4. (Central limit theorem). Let $X_1, X_2, \cdots X_n$ be iid random variables with $E(X_i) = \mu$, $Var(X_i) = \sigma^2 < \infty$.

then.
$$\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \xrightarrow{d} N(0,1).$$

proof of thm4 is NOT required.

. From central limit thm. $\forall\ y$.

$$P\left(\sqrt{n}\left(\frac{\bar{X}_n - \mu}{\sigma}\right) \le y\right) \approx \Phi(y).$$

E.g. 6. If $Y_n \sim$ Exponential (n), $Y=0$, then $Y_n \xrightarrow{d} Y$.

proof: For any $y > 0$.

$$p(Y_n \le y) = \int_0^y n e^{-nx}\,dx = 1 - e^{-ny} \to 1.$$

~~if $y > 0$, then $p(Y_n \le y) \to 0$.~~

For any $y \le 0$, $p(Y_n \le y) = 0$.

since $p(Y \le y) = \begin{cases} 0, & y < 0. \\ 1. & y \ge 0. \end{cases}$ why we don't consider $y = 0$?

b/c $p(Y=0) = 1. \ne 0$.

Ex. 2 ~~Eg. 7~~. If Y_n has ~~den~~ pdf $f_n(y) = \begin{cases} (n+1)\, x^n, & 0 < x < 1 \\ 0, & \text{o/w}. \end{cases}$

~~Then~~. $Y = 1$, then $Y_n \xrightarrow{d} Y$.

proof: For $0 < y \le 1$.

$$p(Y_n \le y) = \int_0^y (n+1)\, x^n\, dx = y^{n+1} \to 0.$$

For $y \le 0$ ~~or $y > 1$~~. $p(Y_n \le y) = 0$.

$y > 1$. $p(Y_n < y) = 1$.

So. $p(Y_n \le y) \to \begin{cases} 0, & y < 1 \\ 1. & y > 1 \end{cases}$

$$p(Y \le y) = \begin{cases} 0, & y < 1. \\ 1. & y > 1. \end{cases}$$

So. $p(Y_n \le y) \to p(Y \le y)$,