Assignment-5

Lanka Prasanna-CS20BTECH11029

Download all latex codes from:

https://github.com/PrasannaLanka/Assignment5/blob/main/Assignment5/codes/Assignment5.tex

PROBLEM UGC/MATH (2018 Dec-Math set-a) Q.104

Let $X_1, X_2, \cdots$ be i.i.d. $N(0,1)$ random variables. Let $S_n = X_1^2+X_2^2+\cdots+X_n^2. \forall n \geq 1$.Which of the following statements are correct?

(A) $\frac{S_n-n}{\sqrt{2}} \sim N(0,1)$ for all $n \geq 1$

(B) For all $\epsilon > 0$,$\Pr\left(\left(\left|\frac{S_n}{n} - 2\right| > \epsilon\right)\right) \to 0$ as $n \to \infty$

(C) $\frac{S_n}{n} \to 1$ with probability 1

(D) $\Pr\left(S_n \leq n + \sqrt{n}x\right) \to \Pr\left(Y \leq x\right) \forall x \in R$,where $Y \sim N(0,2)$

Solution

Theorem 0.1 (chi-square distribution)**.** *If $X_1, X_2, \cdots$ are independent normally distributed random variables with mean 0 and variance 1. Then $\chi = X_1^2 + X_2^2 + \cdots + X_n^2$ is chi-square distributed with n degrees of freedom.*

$$E(\chi) = n \tag{0.0.1}$$

$$Var(\chi) = 2n \tag{0.0.2}$$

Theorem 0.2 (Weak law of large numbers)**.** *Let $X_1, X_2, \cdots$ be i.i.d random variables with same expectation(μ) and finite variance(σ^2).Let $S_n = X_1 + X_2 + \cdots X_n$,Then as $n \to \infty$*

$$\frac{S_n}{n} \xrightarrow{i.p} \mu, \tag{0.0.3}$$

in probability

Proof. Define a new variable

$$X \equiv \frac{X_1 + X_2 + \cdots X_n}{n} \tag{0.0.4}$$

Then, as $n \to \infty$

$$E(X) = E\left(\frac{X_1 + X_2 + \cdots X_n}{n}\right) \tag{0.0.5}$$

$$= \frac{E(X_1) + \cdots E(X_n)}{n} \tag{0.0.6}$$

$$= \frac{n\mu}{n} \tag{0.0.7}$$

$$= \mu \tag{0.0.8}$$

In addition,

$$Var(X) = Var\left(\frac{X_1 + X_2 + \cdots X_n}{n}\right) \tag{0.0.9}$$

$$= Var\left(\frac{X_1}{n}\right) + \cdots Var\left(\frac{X_n}{n}\right) \tag{0.0.10}$$

$$= n\frac{\sigma^2}{n^2} \tag{0.0.11}$$

$$= \frac{\sigma^2}{n} \tag{0.0.12}$$

Therefore, by Chebyshev inequality, for all $\epsilon > 0$,

$$\Pr\left(|X - \mu| \geq \epsilon\right) \leq \frac{Var(X)}{\epsilon^2} = \frac{\sigma^2}{n\epsilon^2} \tag{0.0.13}$$

As $n \to \infty$, it follows that

$$\lim_{n\to\infty} \Pr\left(|X - \mu| \geq \epsilon\right) = 0 \tag{0.0.14}$$

Stated other way as $n \to \infty$

$$\frac{S_n}{n} \xrightarrow{i.p} \mu, \tag{0.0.15}$$

in probability □

Theorem 0.3 (Strong law of large numbers)**.** *Let $X_1, X_2, \cdots$ be i.i.d random variables with same expectation(μ) and finite variance(σ^2).Let $S_n = X_1 + X_2 + \cdots X_n$,Then as $n \to \infty$*

$$\frac{S_n}{n} \xrightarrow{a.s} \mu, \tag{0.0.16}$$

almost surely.

Theorem 0.4 (Central limit theorem)**.** *The Central limit theorem states that the distribution of the sample approximates a normal distribution as the* *sample size becomes larger,given that all the samples are equal in size,regardless of the distribution of the individual samples.*

Definition 0.1 (Almost sure convergence)**.** *A sequence of random variables $\{X_n\}_{n\in N}$ is said to converge almost surely or with probability 1 (denoted by a.s or w.p 1) to X if*

$$\Pr(\omega|X_n(\omega) \to X(\omega)) = 1 \tag{0.0.17}$$

Lemma 0.5 (Properties of mean and variance)**.** *If X is a random variable with a probability density function of $f(x)$. If a and b are constants.*

$$E(X) = \int_R xf(x)dx \tag{0.0.18}$$
$$E(X+Y) = E(X) + E(Y) \tag{0.0.19}$$
$$E(aX+b) = aE(X) + b, \tag{0.0.20}$$
$$Var(X) = E(X^2) - E(X)^2 \tag{0.0.21}$$
$$Var(X+Y) = Var(X) + Var(Y) \tag{0.0.22}$$
$$Var(aX+b) = a^2 Var(X) \tag{0.0.23}$$

Given $X_1, X_2, \cdots$ follow normal distribution with mean 0 and variance 1.

$$f_{X_i}(x) = \frac{1}{\sqrt{2}\pi} e^{-\frac{x^2}{2}}, i \in \{1,2,\cdots\} \tag{0.0.24}$$

(A) From theorem (0.1) S_n is a chi-distributed function with n degrees of freedom.
From (0.0.1)

$$E(S_n) = n \tag{0.0.25}$$

From (0.0.25) and (0.0.20)

$$E\left(\frac{S_n - n}{\sqrt{2}}\right) = \frac{E(S_n) - n}{\sqrt{2}} \tag{0.0.26}$$
$$= \frac{n-n}{\sqrt{2}} \tag{0.0.27}$$
$$= 0 \tag{0.0.28}$$

From (0.0.2)

$$Var(S_n) = 2n \tag{0.0.29}$$

From (0.0.29) and (0.0.23)

$$Var\left(\frac{S_n - n}{\sqrt{2}}\right) = Var\left(\frac{S_n}{\sqrt{2}}\right) \tag{0.0.30}$$
$$= \frac{Var(S_n)}{2} \tag{0.0.31}$$
$$= \frac{2n}{2} \tag{0.0.32}$$
$$= n \tag{0.0.33}$$

Hence,

$$\left(\frac{S_n - n}{\sqrt{2}}\right) \sim N(0,n) \tag{0.0.34}$$

Hence **Option A is false.**

(B) Given

$$S_n = X_1^2 + X_2^2 + \cdots + X_n^2 . \forall n \geq 1 \tag{0.0.35}$$

Assume that For all $\epsilon > 0$,$\Pr\left(\left(\left|\frac{S_n}{n} - 2\right| > \epsilon\right)\right) \to 0$ as $n \to \infty$ is true
From (0.0.25) and (0.0.20)

$$E\left(\frac{S_n}{n}\right) = \frac{E(S_n)}{n} \tag{0.0.36}$$
$$= \frac{n}{n} \tag{0.0.37}$$
$$= 1 \tag{0.0.38}$$

From theorem (0.2)

$$\frac{S_n}{n} \xrightarrow{i.p} E\left(\frac{S_n}{n}\right) \tag{0.0.39}$$

$$\lim_{n\to\infty} \Pr\left(\left|\frac{S_n}{n} - 1\right| > \epsilon\right) = 0 \tag{0.0.40}$$

This means for all $\epsilon > 0$,$\Pr\left(\left|\frac{S_n}{n} - 1\right| > \epsilon\right) \to 0$ as $n \to \infty$
But this is contradiction to our assumption.
Hence **Option B is false .**

(C) Given

$$S_n = X_1^2 + X_2^2 + \cdots + X_n^2 . \forall n \geq 1 \tag{0.0.41}$$

Hence from theorem 0.3 we can write

$$\frac{S_n}{n} \xrightarrow{a.s} E\left(\frac{S_n}{n}\right) \tag{0.0.42}$$

From (0.0.38)

$$\frac{S_n}{n} \xrightarrow{a.s} 1 \tag{0.0.43}$$

almost surely.

$$\Pr\left(\lim_{n\to\infty} \frac{S_n}{n} = 1\right) = 1 \tag{0.0.44}$$

From definition 0.1 and (0.0.44) we can write,

$$\frac{S_n}{n} \xrightarrow{w.p.1} 1 \tag{0.0.45}$$

with probability 1.
Hence **Option C is true**.

(D) Let

$$Y = \frac{S_n - n}{\sqrt{n}} \tag{0.0.46}$$

Using (0.0.25) and (0.0.20)

$$E\left(\frac{S_n - n}{\sqrt{n}}\right) = \frac{E(S_n) - n}{\sqrt{n}} \tag{0.0.47}$$

$$= \frac{n - n}{\sqrt{n}} \tag{0.0.48}$$

$$= 0 \tag{0.0.49}$$

using (0.0.25) and (0.0.20)

$$Var\left(\frac{S_n - n}{\sqrt{n}}\right) = \frac{Var(S_n)}{n} \tag{0.0.50}$$

$$= \frac{2n}{n} \tag{0.0.51}$$

$$= 2 \tag{0.0.52}$$

Hence, from theorem (0.4)

$$Y \sim N(0, 2) \tag{0.0.53}$$

$$\Pr\left(\frac{S_n - n}{\sqrt{n}} \leq x\right) = \Pr\left(S_n \leq n + \sqrt{n}x\right) \tag{0.0.54}$$

Therefore,

$$\Pr\left(S_n \leq n + \sqrt{n}x\right) \to \Pr\left(Y \leq x\right) \forall x \in R \tag{0.0.55}$$

Hence, **Option D is true.**