Chapter Seven

$\boxed{\text{Central limit Theorem}}$

نظرية الغاية المركزية

In this chapter, we shall give some important properties of the theory of large samples in statistical inference.

الاستدلال الاحصائي

متراجحة

7-1 Chebyshev's Inequality

The chebyshev's inequality is a useful theoretical too[l] as well as it helps us to find upper and lower bounds for certain probability when only the me[an] and variance of the probability distribution are k[nown]

Def. 1

If x is a r.v with a finite mean μ and variance σ^2, then for any $k>0$, we have,

$$p\{|x-\mu| \geq k\} \leq \frac{\sigma^2}{k^2} \iff p\{|x-\mu| \geq k\sigma\} \leq \frac{1}{k^2}$$

$$\text{or } p\{|x-\mu| < k\} \geq 1-\frac{\sigma^2}{k^2} \iff p\{|x-\mu| < k\sigma\} \geq 1-$$

28

ex. 1 : Suppose that x is a random variable with mean $\mu = 20$ and variance $\sigma^2 = 16$. Find the probability that x will lie between 15 and 25 ?

Sol.

$$p(15 < x < 25) = p(15-20 < x-\mu < 25-20)$$
$$= p(-5 < x-\mu < 5)$$
$$= p(|x-\mu| < 5)$$

we have, $p(|x-\mu| < k) \geq 1 - \frac{\sigma^2}{k^2}$

$$\therefore \quad p(|x-\mu| < 5) \geq 1 - \frac{16}{25}$$

$$\Longrightarrow p(|x-\mu| < 5) \geq \frac{9}{25}$$

ex. 2 : If it is known that x has a mean of 25 and a variance of 16, then find $p(17 < x < 33)$?

sol. $p(17 < x < 33) = p(17-25 < x-\mu < 33-\mu)$
$= p(-8 < x-\mu < 8) = p(|x-\mu| < 8)$

We have, $p(|x-\mu| < k) \geq 1 - \frac{\sigma^2}{k^2}$

$$\Longrightarrow p(17 < x < 33) = p(|x-\mu| < 8) \geq 1 - \frac{16}{64}$$

7-2 $\boxed{\text{The Weak Law of large Numbers}}$:

This theorem deals with the fact that under some conditions the average of a sequence of random variable converges to the expected average.

Theorem:

let $x_1, x_2, \ldots, x_n$ be a random sample of size n drawn from a population with a finite mean μ.
let $\bar{X}_n$ be the sample mean of the random sample.
Then for any $\epsilon > 0$,

$$p\{|\bar{X}_n - \mu| > \epsilon\} \longrightarrow 0 \quad \text{as} \quad n \longrightarrow \infty$$

7-3 $\boxed{\text{Central Limit Theorem}}$:

Suppose $X_1, X_2, \ldots, X_n$ be a random sample of size n drawn from a population with a finite mean μ and ~~a finite mean μ and~~ a finite variance σ^2. let $\bar{X}_n$ be the sample mean of the random sample.
Consider the normalized random variable:

$$Z_n = \frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{var(\bar{X}_n)}}$$

let the distribution function of Z_n is $F_n(x) = p(Z_n \le x)$,
then, $\lim_{n \to \infty} F_n(x) = \Phi(x)$

توزيع باينوميل

$\boxed{\text{ex } 3}$: let x be the number of successes in 5000 Bernoulli trials with probability of success of 0.7 on a given trial

Use the central limit theorem to estimate :

أي نستخدم بيانات التوزيع الطبيعي

عندما حجم العينة يكبر هذا الشكل توزيع البيانات توزيع طبيعي .

1. $p(x<4950)$.
2. $p(3920<x<4240)$.

sol.

$$p(a<x<b) = p\left(\frac{a-\mu}{\sigma}<\frac{x-\mu}{\sigma}<\frac{b-\mu}{\sigma}\right)$$

$$p(x<a) = p\left(\frac{x-\mu}{\sigma}<\frac{a-\mu}{\sigma}\right)$$

$$\mu = np \Longrightarrow \mu = 5000(0.7) = 3500$$

$$\sigma^2 = npq = (5000)(0.7)(0.3) = 1050$$

$$\sigma = \sqrt{1050} = 32.40$$

$\boxed{1.}$ $p(x<4950) = p\left(\frac{x-\mu}{\sigma}<\frac{4950-3500}{32.40}\right) =$

$= p(z< \quad)$

$= \Phi(\quad) =$

$\boxed{2.}$ $p(3920<x<4240) = p\left(\frac{3920-3500}{32.40}<\frac{x-\mu}{\sigma}<\frac{4240-3500}{32.40}\right)$

$= p(\quad <z< \quad)$

$= \Phi(\quad) - \Phi(\quad)$

5

$e^{-\frac{x-\mu}{\sigma}}$

التوزيع الاسي

2.06

$\boxed{x.4}$: Let $\bar{x}_n = \bar{x}_{15}$ denote the mean of a random of size 15 from the distribution whose p.d.f is :

$\boxed{f(x) = \frac{3}{2}x^2 \quad -1<x<1 \text{, Zero other wise}}$, And $\mu=0, \sigma=$

Use the central limit theorem to Compute :

$p(0.03 \leq \bar{x} \leq 0.15)$?

ex 3

Sol. $p(0.03 \leq \bar{x} \leq 0.15) = p\left(\frac{0.03 - E(\bar{x}_n)}{\sqrt{var(\bar{x}_n)}} \leq \frac{\bar{x} - E(\bar{x}_n)}{\sqrt{var(\bar{x}_n)}} \leq \frac{0.15 - E(\bar{x}_n)}{\sqrt{var(\bar{x}_n)}}\right)$

$E(\bar{x}_n) = E(\bar{x}_5) = \mu = 0$

$var(\bar{x}_n) = var(\bar{x}_{15}) = \frac{\sigma^2}{n} = \frac{\sigma^2}{15} = \frac{3/5}{15} = \frac{3}{75}$

$\therefore p(0.03 \leq \bar{x} \leq 0.15) = p\left(\frac{0.03-0}{\sqrt{\frac{3}{75}}} \leq Z \leq \frac{0.15-0}{\sqrt{\frac{3}{75}}}\right)$

$= p(0.15 \leq Z \leq 0.75)$

$= \Phi(75) - \Phi(0.15)$

$= 0.2138$

$\boxed{\text{ex. 5}}$ If 10 fair dice are rolled, find the approximate probability that the sum obtained is between 30 and 40.
Hint: Use the central limit theorem. Let $\mu_x = 3.5$, $\sigma_x^2 = \frac{35}{12}$

Sol.

$$p\left(30 \leq \sum_{i=1}^{10} x_i \leq 40\right) = ? \qquad \text{let } S = \sum_{i=1}^{10} x_i$$

$$p(30 \leq S \leq 40) = p\left(\frac{30-\mu_s}{\sigma_s} \leq \frac{S-\mu_s}{\sigma_s} \leq \frac{40-\mu_s}{\sigma_s}\right)$$

$$\mu_s = E(S) = E\sum_{i=1}^{10} x = \sum_{i=1}^{10} E(x)$$

$$= \sum_{i=1}^{10} 3.5 = (10)(3.5) = 35$$

$$\sigma_s^2 = var(S) = var\sum_{i=1}^{10} x_i$$

$$= \sum_{i=1}^{10} var(x_i) = \sum_{i=1}^{10} \frac{35}{12} = (10)\frac{35}{12} = \frac{350}{12}$$

$$\therefore\ p(30 \leq S \leq 40) = p\left(\frac{30-35}{\sqrt{\frac{350}{12}}} \leq Z \leq \frac{40-35}{\sqrt{\frac{350}{12}}}\right)$$

$$= p(-0.9257 \leq Z \leq 0.9257)$$

$$= \Phi(0.9257) - \Phi(-0.9257)$$

$$= 0.65$$

$\boxed{\text{ex. 7}}$: Suppose 3% of the items made by a factory are ... A sample of 100 items is drawn at random. What is ... that it will contain exactly 2 defective items? Using:

a. Binomial distribution?

b. poisson distribution?

Sol. $n=100$, $p=\frac{3}{100}$, $q=1-p=\frac{97}{100}$

$\boxed{a.}$

$$p(x)=\begin{cases} C_x^n \, p^x \, q^{n-x} & x=0,1,2,\ldots,n \\ 0 & \text{o.w} \end{cases}$$

$$p(x)=\begin{cases} C_x^{100} \, (0.03)^x \, (0.97)^{100-x} & x=0,1,2,\ldots,100 \\ 0 & \text{o.w} \end{cases}$$

$$\Longrightarrow p(x=2)=C_2^{100} \, (0.03)^2 \, (0.97)^{100-2} = 0.2251$$

$\boxed{b.}$

$$p(x)=\begin{cases} \dfrac{e^{-\lambda} \lambda^x}{x!} & x=0,1,2,\ldots \\ 0 & \text{o.w} \end{cases}$$

$\lambda=np=100(0.03)=3$

$$p(x)=\begin{cases} \dfrac{e^{-3} 3^x}{x!} & x=0,1,2,\ldots \\ 0 & \text{o.w} \end{cases}$$

$$\Longrightarrow p(x=2)=\frac{e^{-3} 3^2}{2!} = 0.2240$$

The difference is 0.0011

7.5 The Hypergeometric Approximation to the Binomial distribution :

Theorem : Suppose that x is a hypergeometric random variable with parameters , n, N, D . let $p = \frac{D}{N}$, if D and N-D are large , n and x are small compared with D and N-D, then the p.m.f of x given by :

$$p(x) = \begin{cases} C_x^n \, p^x q^{n-x} & x = 0, 1, 2, \ldots, n \\ 0 & o.w \end{cases} \qquad p = \frac{D}{N} .$$

$\boxed{ex. 8}$

let x be random variable such that $x \sim Hyp(20, 200, 1000)$

If x=20 , find p(x=20) ? Use binomial dist. ?

sol. $p = \frac{D}{N} = \frac{200}{1000} = 0.2 \Longrightarrow q = 0.8$

$$p(x) = \begin{cases} C_x^{20} (0.2)^x (0.8)^{20-x} & x = 0, 1, \ldots, 20 \\ 0 & o.w \end{cases}$$

$$p(x=20) = C_{10}^{20} (0.2)^{10} (0.8)^{20-10} = 0.0020$$

problems

1. If x is a random variable such that $E(x)=3$, and $E(x)^2=13$. Use chebyshev's inequality to determine a lower bound for the probability $p(-2<x<8)$?

2. If x is a random variable with mean 33 and variance 16, use chebyshev's inequality to determine a lower bound for $p(23<x<43)$?

3. Suppose that x is a random variable with mean and variance both equal to 20. What can be said about $p(0\leq x\leq 40)$?

4. let $x_1, x_2, \ldots, x_{20}$ be independent poisson random variables with mean 1. Use the central limit theorem to approximate $p(\sum_{i=1}^{20} x_i > 15)$?

5. let $\bar{x}_{36}$ be the mean of a random sample of size 36 from an exponential distribution with mean 3. Approximate $p(2.5<\bar{x}_{36}\leq 4)$?

6. let $\bar{x}_{12}$ be the mean of a random sample of size 12 from an uniform distribution on the interval $(0,1)$. Approximate $p(\frac{1}{2}\leq \bar{x}_{12}\leq \frac{2}{3})$?

$\boxed{7.}$ Approximate $p(39.5 \leq \bar{X}_{32} \leq 41.25)$, where $\bar{X}_{32}$ is the mean of a random sample of size 32 from a distribution with mean $\mu = 40$ and variance $\sigma^2 = 8$.

$\boxed{8.}$ let $Y_{15} = X_1 + X_2 + \cdots + X_{15}$ be the sum of a random sample of size 15 from the distribution whose p.d.f is,

$$f(x) = \begin{cases} \frac{3}{2}x^2 & -1 < x < 1 \\ 0 & \text{o.w} \end{cases}$$

Approximate $p(-0.3 \leq Y_{15} \leq 1.5)$?

Centeral limit theorem

for $\bar{x}$ we have.

$E\bar{x} = \mu$, $V(\bar{x}) = \frac{\sigma^2}{n}$ 2.35

Thus as n increasing the variance of $\bar{x}$ decrease.

The dist. of $\bar{x}$ clearly depend on n.

so we are dealing with sequence of dist.

To denote this we will place the subscipt n on the random Sample of size n, we write $\bar{X}_n$ in the case of the mean of a random Sample of size n.

By using the chebyshev's inequality consider $\epsilon > 0$

$$P\{|\bar{x}_n - \mu| \geq \epsilon\}$$

the random sample of size n , we write $\bar{X}_n$

· mean of a random sample of size n.

By using the chebyshev's inequality consider $\epsilon > 0$

$$P\{|\bar{x}_n - \mu| \geq \epsilon\}$$

$$= P\left\{|\bar{x}_n - \mu| \geq \left(\frac{\epsilon\sqrt{n}}{\sigma}\right)\left(\frac{\sigma}{\sqrt{n}}\right)\right\}$$

But the standard deviation of $\bar{x}$ is $\frac{\sigma}{\sqrt{n}}$

Hence using chebyshev's inequality with $k = \frac{\epsilon\sqrt{n}}{\sigma}$

~~where~~ we have

$$P\{|\bar{x}_n - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2\sqrt{n}}$$

Since the prob. is non negative it follows that.

$$\lim P\{|\bar{x}_n - \mu| \geq \epsilon\} \leq \lim_{n\to\infty} \frac{\sigma^2}{\epsilon^2 n} = 0$$

this implies that.

Hence using chebyshev's inequality with $K = \frac{\epsilon \sqrt{n}}{\sigma}$

~~where~~ we have

$$p\{|\bar{x}_n - \mu| \geq \epsilon\} \leq \frac{\sigma^2}{\epsilon^2 \sqrt{n}}$$

since the prob. is non negative it follows that.

$$\lim p\{|\bar{x}_n - \mu| \geq \epsilon\} \leq \lim_{n\to\infty} \frac{\sigma^2}{\epsilon^2 n} = 0$$

this implies that.

$$\lim p\{|\bar{x}_n - \mu| \geq \epsilon\} = 0$$

or equivalently

$$\lim p[|\bar{x}_n - \mu| < \epsilon] = 1$$

38

الرجيد للمقايس

The preceding discussion show that the prob. associated with the dist. of $\bar{x}_n$ becomes concentrated in an arbitrary small interval about μ as n increases. This is one of the form of the law of large numbers

P. 154

Example Let $x_1, x_2, \dots, x_n$ denote a r.s. from a dist. $b(1,p)$, $0<p<1$

~~this~~ we know that.

$$Y_n = x_1 + x_2 + \dots + x_n.$$

from the preceding results with ~~[illegible]~~

show that.

$$\lim P\left[\left|\frac{Y_n}{n} - p\right| < \epsilon\right]$$

Theorem Central limit theorem

Let $\bar{X}_n$ be the mean of a r. s. $X_1, \ldots X_n$ of size n from a dist. with a finite mean μ and a finite positive variance σ^2, then the dist of.

$$W_n = \frac{\bar{X}_n - \mu}{\sigma / \sqrt{n}} = \frac{\sum_{i=1}^{n} X_i - n\mu}{\sqrt{n}\,\sigma}$$

is $N(0,1)$ in the limit as $n \to \infty$

١٠/٤/٢٠١٩ الأربعاء
2.45

The central limit theorem

or Convergence to Normal dist (Lindeberg's theorem)

let $X_1, X_2, \ldots$ be indep. r.v.'s with each X_k having finite mean m_k and finite variance σ_k^2

let $S_n = \sum_{k=1}^{n} X_k \quad , k = 1, 2, \ldots$

then

$$E(S_n) = \sum_{k=1}^{n} m_k$$

$$Var(S_n) = C_k^2 = \sum_{k=1}^{n} \sigma_k^2$$

اذا كان لدينا $X_1, X_2, \ldots$ متغيرات عشوائية مستقلة ومع كل X_k ~~...~~

لدينا متوسط m_k محدد

وتباين σ_k^2 محدد.

$\underbrace{[X_1 \ldots]}_{X_k} \underbrace{[X_1 \ldots X]}_{X_k} \ldots$

لدينا k من المجاميع الجزئية n.

وان لكل k متوسط محدد m_k وتباين محدد σ_k^2

وعلى هذا نعرف

$S_n = \sum_{k=1}^{n} X_k \quad , n = 1, 2, \ldots$

$E(S_n)$ المتوسط

$V(S_n)$ التباين

نظرية النهاية المركزية

$X \sim Bernoulli\ (p)$

$EX = p$

$V(X) = pq$

$n \to \infty$

$X \sim N(m = p, \sigma^2 = pq)$

230

41

let $T_n = C_n^{-1} \left(S_n - E(S_n) \right)$ standraization

F_K be the dist. fun. of X_K

if for $\epsilon > 0$

$$\frac{1}{C_n^2} \sum_{K=1}^{n} \int_{\{x : |X - m_K| \geq \epsilon C_n\}} (X - m_K)^2 \, dF_K(x) \longrightarrow 0 \quad \text{as } n \to \infty$$

chebyshevs

Then T_n converges in dist. to normal variable X^*

i.e $N(0,1)$

المؤدي لك

Implications of lindeberg's theorem

لو اعطي هذا البديله .

① the Uniformly bounded case

Assume $|X_k| \leq M \quad \forall k$ and $C_n \to \infty$

then
$$\int_{\{x:|x-m_k| \geq \epsilon C_n\}} (x-m_k)^2 \, dF_k(x) = E(X_k - m_k)^2 \, I_{\{|X_k - m_k| \geq}$$

لو كان تقسم على الـ

by chebyshev's

لو اخذنا الجمع sum وقسمنا على C_n^2

By chebyshev's ine

$$\therefore \frac{1}{C_n^2} \sum_{k=1}^{n} \int_{\{x:|x-m_k| \leq \epsilon C_n\}} (x-m_k)^2 \, dF_k(x) \leq (2M)^2 \, \underbrace{P\{|x-m_k| \geq \epsilon C}_{\leq \frac{1}{\epsilon^2} \frac{\sigma_k^2}{C_n^2}}$$

$$\frac{1}{C_n^2} \sum_{k=1}^{n} \int_{\{x:|x-m_k| \leq \epsilon C_n\}} (x-m_k)^2 \, dF_k(x) \leq \frac{2M^2}{\epsilon^2 C_n^2} \cdot \sigma_k^2 \longrightarrow 0$$

as $C_n \to \infty$

بشرط .

متطابق

The identical dist. case

Assume that the X_k are i.i.d, with finite mean M and finite variance $\sigma^2 > 0$

if F is the dist. of the X_k, then

$$\frac{1}{C_n^2} \sum_{k=1}^{n} \int_{\{x: |x-m_k| > \epsilon C_n\}} (x-m_k)^2 \, dF_k(x)$$

$$= \frac{1}{n\sigma^2} \sum_{k=1}^{n} \int_{\{x: |x-m_k| \geq \epsilon\sigma\sqrt{n}\}} (x-m_k)^2 \, dF(x)$$

$$= \frac{1}{\sigma^2} \int_{\{x: |x-m| \geq \epsilon\sigma\sqrt{n}\}} (x-m)^2 \, dF(x) \longrightarrow 0$$

since σ^2 is finite $\{x: |x-m| \geq \epsilon\sigma\sqrt{n}\} \downarrow \emptyset$ as $n \to \infty$

هذا ال set بيصير ال $\emptyset$.

The Bernoulli case

let S_n be the number of successes in n Bernoulli trials with prob. of successes p on a given trial we may write

$$S_n = X_1 + X_2 + \cdots + X_n$$

where the X_k are indep. and

$$P\{X_k = 1\} = p$$

$$P\{X_k = 0\} = 1 - p = q$$

we may take X_k as the indecator of a success on trial k.

Thus Case 2 applies with

$$m = E(X_k) = p$$

$$\sigma^2 = E(X_k^2) - (E X_k)^2 = npq.$$

$$E(S_n) = nm = np$$

$$C_n^2 = n\sigma^2 = npq$$

متغير عشوائي يتوزع طبيعي

Thus $T_n = \dfrac{S_n - np}{(npq)^{1/2}}$ and $T_n \xrightarrow{d} X^*$

That is $P\{T_n \leq x\} \to F^*(x)$, $\forall x$ where

$$F^*(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{1}{2}t^2}\, dt$$ standard normal

Liapounov's Condition

Assume that.

$$\frac{1}{C_n^{2+\delta}} \sum_{k=1}^{n} E\left[|X_k - m_k|^{2+\delta}\right] \longrightarrow 0$$

for som $\delta > 0$

$$E\left[|X_k - m_k|^{2+\delta}\right] = \int_{-\infty}^{\infty} |x - m_k|^{2+\delta} \, d f_k(x)$$

$$\geq \int_{\{x : |x - m_k| \geq \epsilon C_n\}} |x - m_k|^{\delta} \cdot |x - m_k|^{2} \, d f_k(x)$$

$$\geq \epsilon^{\delta} C_n^{\delta} \int_{\{x : |x - m_k| \geq \epsilon C_n\}} (x - m_k)^2 \, d f_k(x)$$

then $$\frac{1}{C_n^2} \sum_{k=1}^{n} \int_{\{x : |x - m_k| \geq \epsilon C_n\}} (x - m_k)^2 \, d f_k(x)$$

$$\leq \frac{1}{C_n^2} \sum_{k=1}^{n} \frac{E\left[|X_k - m_k|^{2+\delta}\right]}{\epsilon^{\delta} C_n^{\delta}} \longleftarrow$$ واجب توضيح

$$= \sum_{k=1}^{n} \frac{E\left[|X_k - m_k|^{2+\delta}\right]}{\epsilon^{\delta} C_n^{2+\delta}} \longrightarrow 0$$

الرط

Martingals

حاول
محفز

(Cp4 a)

Def. let (Ω, F, P) be a prob. space $\{X_1, X_2, \dots\}$ a sequence of integrable r.v.'s on (Ω, F, P) and $F_1 \subset F_2 \subset \dots$ an increasing sequence of sub σ-field on F.

X_n is assumed F_n measurable that is

i.e $X_n : (\Omega, F_n) \longrightarrow (R, B(R))$

The sequence $\{X_n\}$ is said to be martingle relative to F_n iff:

قريب

for all $n = 1, 2, \dots$

① $E(X_{n+1} / F_n) = X_n$ a.e (all most every where)

② a sub-martingals iff

$$E(X_{n+1}/F_n) \geq X_n \quad \text{a.e} \quad \text{and.}$$

③ a super martingals iff:

$$E(X_{n+1}/F_n) \leq X_n \quad , \text{a.e}$$

علمًا بأن X_{n+1} معلوم في F_{n+1}

Comments

(a) if $\{X_n, F_n\}$ is amartingals then

$$E(X_{n+k}/F_n) = X_n \qquad n, k = 1, 2, 3, \ldots$$

i.e

$$\begin{aligned} E(X_{n+2}/F_n) &= E\left[E(X_{n+2}/F_{n+1})/F_n\right] \\ &= E(X_{n+1}/F_n) \\ &= X_n \end{aligned}$$

(b) if $\{X_n, F_n\}$ is amartingale, then

$$E(X_{n+1}/X_1, \ldots, X_n) = X_n \qquad , n = 1, 2, \ldots$$

(c) if $\{X_n, F_n\}$ is amartingale iff:

$$\int_A X_n \, dp = \int_A X_{n+1} \, dp \quad \forall A \in F_n \qquad , n = 1, 2, \ldots$$

let $Y_1, Y_2, \ldots$ be independent r.v.'s with zero mean and set $X_n = \sum_{k=1}^{n} Y_k$

$F_n = F(Y_1, \ldots, Y_n)$ then $\{X_n, F_n\}$ is a martingale

<u>sol</u>

$$E(X_{n+1}/F_n) = E(X_n + Y_{n+1}/\, Y_1, Y_2, \ldots, Y_n)$$

$$= X_n + E(Y_{n+1}/Y_1, Y_2, \ldots, Y_n)$$

since X_n is F_n – measurable

$$\Rightarrow \quad = X_n + E(Y_{n+1}) \quad \text{By independency}$$

$$= X_n + 0$$

$$= X_n$$

let $Y_1, Y_2, \cdots$ be indep. r.v.'s with

$E(y_j) = a_j \neq 0$

and set $X_n = \prod_{j=1}^{n} \left(\frac{Y_j}{a_j} \right)$,

$F_n = f(Y_1, \ldots, Y_n)$ then $\{X_n, F_n\}$ is a martingle

sol

$$E(X_{n+1} / F_n) = E\left[\left(\frac{X_n Y_{n+1}}{a_n}\right) / Y_1, Y_2, \cdots Y_n\right]$$

$$= X_n E\left(\frac{Y_{n+1}}{a_{n+1}}\right)$$

$$= X_n \left(\frac{a_{n+1}}{a_{n+1}}\right)$$

$$= X_n$$

فيبي
P261–262

Central limit theorem

Consider a sequence of independent and identically dist. random variables. $\{X_n\}$ with mean $E(X_n) = \mu$ and $Var\{X_n\} = \sigma^2$

let

$$S_n = \sum_{i=1}^{n} x_i \quad , \quad \bar{X}_n = \frac{S_n}{n}$$

Then

$$E(\bar{X}_n) = E\,\frac{(X_1 + \cdots + X_n)}{n} = \frac{n\,\mu}{n} = \boxed{\mu}.$$

and

$$Var(\bar{X}_n) = Var\left(\frac{X_1 + \cdots + X_n}{n}\right) = \frac{n\,\sigma^2}{n^2} = \boxed{\frac{\sigma^2}{n}}$$

it is clear that $Var(\bar{X}_n) \longrightarrow 0$ as $n \longrightarrow \infty$

This means that the density fun. of $\bar{X}_n$ is Concentrated around the mean μ for large n. we shall see later that as $n \longrightarrow \infty$ the dist. fun. of $\bar{X}_n$ tends to the normal dist. under certain Conditions. This property is one of the forms of the Centeral limit theorem which is.

263 ص

Theorem

let $x_1, x_2, \ldots$ be a sequence of i.i.d r.v.'s each with mean μ and variance σ^2, then the dist. of $Z_n = \dfrac{S_n - E(S_n)}{\sqrt{var(S_n)}}$ tends to the standard normal as $n \to \infty$, that is

$$P\{Z_n \le z\} = \phi(z) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \, dt \quad \text{as } n \to \infty$$

البرهان واجب
وهو سؤال معظّم في
الامتحان النهائي
أما التدريب .

proof: the theorem of Convergence in dist. that is.

The Theorem

> let $\phi_n(t)$ be the characteristic fun. of X_n. if $X_n \xrightarrow{d} X$ then $\phi_n(t) \to \phi(t)$, where $\phi(t)$ is the c.d.f. of X.
> if $\phi_n(t) \longrightarrow \phi(t)$ and the limit fun. is continuous at $t = 0$, then
> $$X_n \xrightarrow{d} X$$

نص البرهان

it is sufficient to show that $\phi_{Z_n}(t) \longrightarrow \phi(t)$ as $n \to \infty$ where $\phi_{Z_n}(t)$ and $\phi(t)$ are the c.d.f. of Z_n and Z (standard Normal dist.). respectively. we have.

P263

$E(S_n) = n\,E(X_i) = n\mu$

and $Var(S_n) = n\,Var(X_i) = n\sigma^2$

also we have.

$$\phi_{S_n}(t) = \phi_{X_1+\cdots+X_n}(t) = \prod_{i=1}^{n} \phi_{X_i}(t)$$

$$= [\phi_x(t)]^n$$

beause X_i are i.i.d. r.v.'s

There fore.

$$Z_n = \frac{S_n - n\mu}{\sqrt{n}\,\sigma}$$

There fore

$$\phi_{Z_n}(t) = e^{\frac{-itn\mu}{\sqrt{n}\sigma}} \cdot \phi_{S_n}\left(\frac{1}{\sqrt{n}\sigma}\right)^n$$

$$= e^{\frac{-it\sqrt{n}\mu}{\sigma}} \left[\phi_x\left(\frac{1}{\sqrt{n}\sigma}\right)\right]^n$$

By the taylor series expansion of $\phi_x(t)$ we get.

$$\phi_x(t) = 1 + itE(x) - \frac{t^2}{2!}E(x^2) + o(t^2)$$

There fore

$$\phi_x\left(\frac{1}{\sqrt{n}\sigma}\right) = 1 + \frac{it}{\sqrt{n}\sigma}E(x) - \frac{t^2}{2n\sigma^2}E(x^2) + o\left(\frac{t^2}{n\sigma^2}\right)$$

and

P263

$E(S_n) = n\,E(X_i) = n\mu$

and $Var(S_n) = n\,Var(X_i) = n\sigma^2$

also we have.

$$\phi_{S_n}(t) = \phi_{X_1+\dots+X_n}(t) = \prod_{i=1}^{n} \phi_{X_i}(t)$$

$$= [\phi_x(t)]^n$$

beause X_i are i.i.d. r.v.'s

There fore.

$$Z_n = \frac{S_n - n\mu}{\sqrt{n}\,\sigma}$$

There fore

$$\phi_{Z_n}(t) = e^{\frac{-itn\mu}{\sqrt{n}\sigma}} \cdot \phi_{S_n}\left(\frac{1}{\sqrt{n}\,\sigma}\right)^n$$

$$= e^{\frac{-it\sqrt{n}\,\mu}{\sigma}} \left[\phi_x\left(\frac{1}{\sqrt{n}\,\sigma}\right)\right]^n$$

By the taylor series expansion of $\phi_x(t)$

we get.

$$\phi_x(t) = 1 + itE(x) - \frac{t^2}{2!}E(x^2) + o(t^2)$$

There fore

$$\phi_x\left(\frac{1}{\sqrt{n}\,\sigma}\right) = 1 + \frac{it}{\sqrt{n}\,\sigma}E(x) - \frac{t^2}{2n\sigma^2}E(x^2) + o\left(\frac{t^2}{n\sigma^2}\right)$$

and

$$\phi_{z_n}(t) = e^{\frac{-it\sqrt{n}\,\mu}{\sigma}} \cdot \left[1 + \frac{it\,E(x)}{\sqrt{n}\,\sigma} - \frac{t^2}{2n\sigma^2} E(x^2) + o\left(\frac{t^2}{n\sigma^2}\right)\right]^n$$

$$= e^{\frac{-it\sqrt{n}}{\sigma}} \left[1 + \frac{it\,\mu}{\sqrt{n}\,\sigma} - \frac{t^2}{2n\sigma^2}(\sigma^2 + \mu^2) + o\left(\frac{t^2}{n\sigma^2}\right)\right]^n$$

where $Ex = \mu$, $V(x) = \sigma^2 = Ex^2 - (Ex)^2$

$\sigma^2 = Ex^2 - \mu^2$

$$\Rightarrow \quad = e^{\frac{-it\sqrt{n}}{\sigma}} \left[1 + \frac{it\mu}{\sqrt{n}\,\sigma} - \frac{t^2}{2n} + o\left(\frac{t^2}{n\sigma^2}\right)\right]^n$$

but we have.

$$\left[1 + \frac{\frac{it\mu}{\sqrt{n}\,\sigma} - \frac{t^2}{2}}{n} + o\left(\frac{t^2}{n\sigma^2}\right)\right]^n$$

$$\longrightarrow e^{\frac{it\mu\sqrt{n}}{\sigma} - \frac{t^2}{2}} \quad \text{as } n \longrightarrow \infty$$

therefore.

$$\lim_{n\to\infty} \phi_{z_n}(t) = e^{-\frac{it\sqrt{n}\,\mu}{\sigma}} \cdot e^{\frac{it\mu\sqrt{n}}{\sigma} - \frac{t^2}{2}}$$

$$= e^{-\frac{t^2}{2}}$$

$$= \phi_z(t)$$

where $\phi_z(t) = e^{-\frac{t^2}{2}}$ is the c.d.f. of the standard Normal dist.

Therefor the dist. of Z_n approaches the standard Normal distribution.