Chapter 5. part 2.

§5.3 Convergence in distribution, Central Limit Theorem.

Definition. Let $X_1, X_2, \cdots$ be a sequence of r.v's with CDF $F_1, F_2, \cdots, F_n, \cdots$ and let X be a r.v. with CDF, F, i.e. $F(x) = P(X \le x)$, $F_n(x) = P(X_n \le x)$.

$X_n \longrightarrow X$ in distribution, if $\lim_{n\to\infty} F_n(x) = F(x)$, at every continuous point of $F(x)$.

From a fact: distribution of F_n is determined by its Moment generate function (m.g.f). $M_n(t)$, $M_n(t) = E(e^{tX_n})$, one has

Theorem A. Continuity Theorem.

Let F_n, a sequence of CDF with corresponding M_n. Let F be a CDF with m.g.f. M. If $M_n(t) \longrightarrow M(t)$ for all t in an open interval containing zero, then $F_n(t) \longrightarrow F(t)$ for all continuity points of F.

Example A. Poisson distribution can be approximated by normal distribution for large λ.

i.e. $X_n \sim P(\lambda_n)$, $P(X_n = k) = \frac{\lambda_n^k}{k!} e^{-\lambda_n}$, $k = 0, 1, 2, \cdots$ $\quad E(X_n) = \lambda_n$, $Var(X_n) = \lambda_n$.

when $\lambda_n \to \infty$, $X_n \stackrel{\cdot}{\sim} N(\lambda_n, \lambda_n)$

or $\frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \stackrel{\cdot}{\sim} N(0,1)$. R-demonstration

Next, we show that $Z_n = \frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \xrightarrow{d} N(0,1)$ as $\lambda_n \to \infty$, using Thm A.

idea is to show $M_{Z_n}(t) = E(e^{tZ_n}) \longrightarrow E(e^{tZ}) = e^{t^2/2}$, where $Z = N(0,1)$.

Proof: Write $Z_n = \frac{1}{\sqrt{\lambda_n}} X_n - \sqrt{\lambda_n}$,

thus, $M_{Z_n}(t) = e^{-\sqrt{\lambda_n}\, t} M_{X_n}\left(\frac{1}{\sqrt{\lambda_n}} t\right)$

$\left(\begin{array}{l} \text{using } Y = a + bX, \\ M_Y(t) = e^{at} M_X(bt) \end{array}\right)$

$Z_n = \frac{1}{\sqrt{\lambda_n}} X_n - \sqrt{\lambda_n}$, $a = -\sqrt{\lambda_n}$, $b = \frac{1}{\sqrt{\lambda_n}}$

Since $X_n \sim P(\lambda_n)$, $M_{X_n}(t) = E(e^{tX_n}) = e^{\lambda_n(e^t - 1)}$ from Chapter 4.

$\longrightarrow M_{X_n}\left(\frac{t}{\sqrt{\lambda_n}}\right) = e^{\lambda_n\left(e^{\frac{t}{\sqrt{\lambda_n}}} - 1\right)}$.

One has $M_{Z_n}(t) = e^{-\sqrt{\lambda_n}\, t} \cdot e^{\lambda_n\left(e^{\frac{t}{\sqrt{\lambda_n}}} - 1\right)}$,

$= e^{-\sqrt{\lambda_n}\, t + \lambda_n\left(e^{\frac{t}{\sqrt{\lambda_n}}} - 1\right)}$.

We want to show $M_{Z_n}(t) \to e^{t^2/2}$, which is equivalent to show $\log M_{Z_n}(t) \to \frac{t^2}{2}$

or $(*)$ $\quad -\sqrt{\lambda_n}\, t + \lambda_n\left(e^{\frac{t}{\sqrt{\lambda_n}}} - 1\right) \longrightarrow t^2/2$, as $\lambda_n \to \infty$.

same as $\quad \lambda_n\left(e^{\frac{t}{\sqrt{\lambda_n}}} - 1 - \frac{t}{\sqrt{\lambda_n}}\right) \longrightarrow \frac{t^2}{2}$.

let $u = \frac{1}{\sqrt{\lambda_n}} \to 0$, $\quad \lim_{\lambda_n \to \infty} \lambda_n\left(e^{\frac{t}{\sqrt{\lambda_n}}} - 1 - \frac{t}{\sqrt{\lambda_n}}\right) = \lim_{u \to 0} \frac{e^{tu} - 1 - tu}{u^2} = \lim_{u \to 0} \frac{te^{tu} - 0 - t}{2u}$

(or $\lambda_n = \frac{1}{u^2} \to \infty$)

$$= \lim_{u \to 0} \frac{te^{tu} - t}{2u} = \lim_{u \to 0} \frac{t^2 e^{tu} - 0}{2} = \frac{t^2}{2}.$$

Thus, we proved that $M_{Z_n}(t) \longrightarrow M_Z(t)$, as $\lambda_n \to \infty$.

From Thm A, one has $\quad Z_n = \frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \longrightarrow Z = N(0,1)$ in distribution.

Therefore $\quad P(Z_n \le x) \longrightarrow P(Z \le x)$

or $\quad P\left(\frac{X_n - \lambda_n}{\sqrt{\lambda_n}} \le x\right) = P(X_n \le \lambda_n + \sqrt{\lambda_n} \cdot x) \longrightarrow P(Z \le x) = \Phi(x)$.

Example B. $\quad X \sim P(100)$, i.e. $P(X = k) = \frac{100^k}{k!} e^{-100}$. $\lambda = 100$.

from above discussion, $X \mathrel{\dot\sim} N(100, 100) = N(100, 10^2)$.

$P(X \le 110) = \sum_{k=0}^{110} \frac{100^k}{k!} e^{-100} = 0.8528627$ (= ppois(110, 100) in R)

using normal approximation. $P(X \le 110) \approx P(W \le 110)$ where $W \sim N(100, 10^2)$

= pnorm(110, 100, 10) = 0.8413447. (compare exact value 0.8528627).

Central Limit Theorem

$X_1, X_2, \cdots$ is a sequence of i.i.d rv's with $\mu = E(X_1)$ and $\sigma^2 = Var(X_1)$.

Let $S_n = X_1 + X_2 + \cdots + X_n$, then $E(S_n) = n \cdot \mu$, $Var(S_n) = n \cdot \sigma^2$.

let $$Z_n = \frac{S_n - E(S_n)}{\sqrt{Var(S_n)}} = \frac{S_n - n\mu}{\sqrt{n \cdot \sigma^2}} = \frac{S_n - n \cdot \mu}{\sqrt{n}\,\sigma}.$$

Theorem B. Central Limit Theorem (CLT).

$$Z_n = \frac{S_n - n \cdot \mu}{\sqrt{n}\,\sigma} \longrightarrow N(0,1) = Z \text{ in distribution,}$$

or $$P(Z_n \le x) = P\left(\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \le x\right) \longrightarrow P(Z \le x) = \Phi(x).$$

An application. Let $X_1, X_2, \cdots X_{10}$ be i.i.d. $\sim U(0,1)$. Calculate an approximation to $P(\sum_{i=1}^{10} X_i > 6)$.

Ans. $S_{10} = \sum_{i=1}^{10} X_i$, from $E(X_1) = 0.5$, $Var(X_1) = \frac{1}{12}$, one has $\mu = 0.5$, $\sigma^2 = \frac{1}{12}$, $\sigma = 1/\sqrt{12}$.

From CLT. $$P\left(\sum_{i=1}^{10} X_i > 6\right) = P(S_{10} > 6) = P\left(\frac{S_{10} - 10 \cdot (0.5)}{\sqrt{10} \cdot 1/\sqrt{12}} > \frac{6 - 10 \cdot (0.5)}{\sqrt{10} \cdot 1/\sqrt{12}}\right)$$

$$\doteq P\left(Z > \frac{6-5}{\sqrt{10/12}}\right) = P(Z > \sqrt{1.2}) = 1 - \Phi(\sqrt{1.2}) = 0.1367.$$

Proof: $$Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}} = \frac{\sum_{i=1}^{n}(X_i - \mu)}{\sqrt{n \cdot Var(X_i - \mu)}} = \frac{\sum_{i=1}^{n} X_i'}{\sqrt{n \cdot Var(X_i')}},$$ where $X_i' = X_i - \mu$, $E(X_i') = 0$

(note $Var(X_i) = Var(X_i - \mu)$)

thus, we only provide a proof for CLT with $EX_1 = 0$.

i.e. we show $Z_n = \frac{S_n}{\sqrt{n}\,\sigma} \longrightarrow N(0,1)$ in distribution, where $X_1, \cdots, X_n$ iid, $E(X_1) = 0$, $Var(X_1) = \sigma^2$.

The idea is to show $M_{Z_n}(t) \longrightarrow M_Z(t) = e^{t^2/2}$, where $M_{Z_n}(t) = E(e^{tZ_n}) =$

$$= E\left(e^{t\frac{S_n}{\sqrt{n}\sigma}}\right) = E\left(e^{t\cdot\frac{\sum X_i}{\sqrt{n}\sigma}}\right) = E\left(e^{\sum_{i=1}^{n}\frac{tX_i}{\sqrt{n}\sigma}}\right) = E\left(\prod_{i=1}^{n} e^{\frac{tX_i}{\sqrt{n}\sigma}}\right)$$

$$\underset{\text{independent}}{=} \prod_{i=1}^{n} E\left(e^{\frac{tX_i}{\sqrt{n}\sigma}}\right) \underset{\text{distribution}}{\overset{\text{same}}{=}} \left[E\left(e^{\frac{t}{\sqrt{n}\sigma}X_1}\right)\right]^n. \qquad (**)$$

Now, let $M_{X_1}(t) = E(e^{tX_1})$, m.g.f. for r.v. X_1. $E(X_1) = M'_{X_1}(0)$, from $E(X_1) = 0$, one has $M'_{X_1}(0) = 0$. From $E(X_1^2) = M''_{X_1}(0)$, one has $M''_{X_1}(0) = Var(X_1) = \sigma^2$.

and $Var(X_1) = E(X_1^2) - (EX_1)^2 = E(X_1^2) = \sigma^2$, $EX_1 = 0$.

Now, $M_{X_1}(s) = M_{X_1}(0) + M'_{X_1}(0)(s-0) + \frac{M''_{X_1}(0)}{2}(s-0)^2 + \varepsilon_s$ (Taylor expansion for $M_{X_1}(s)$ around $s_0 = 0$.)

$$= 1 + 0\cdot s + \frac{\sigma^2}{2}s^2 + \varepsilon_s,$$

$$= 1 + \frac{\sigma^2 s^2}{2} + \varepsilon_s.$$

(where $M_{X_1}(0) = E(e^{0\cdot X_1}) \equiv 1$, and $\frac{\varepsilon_s}{s^2} \longrightarrow 0$, as $s \to 0$.)

From above $(**)$, one has $M_{Z_n}(t) = \left[E\left(e^{\frac{t}{\sqrt{n}\sigma}X_1}\right)\right]^n$

where $E\left(e^{\frac{t}{\sqrt{n}\sigma}X_1}\right) = E(e^{sX_1}) = M_{X_1}(s)$, with $s = \frac{t}{\sqrt{n}\sigma} \longrightarrow 0$ as $n \to \infty$.

thus $M_{Z_n}(t) = [M_{X_1}(s)]^n = \left[1 + \frac{\sigma^2}{2}\frac{t^2}{n\sigma^2} + \varepsilon_n\right]^n,$

$$= \left[1 + \frac{1}{n}\left(\frac{t^2}{2} + n\varepsilon_n\right)\right]^n.$$

where $\frac{\varepsilon_n}{\frac{t^2}{n\sigma^2}} \longrightarrow 0$, as $n \to \infty$

or $\frac{\sigma^2}{t^2} n\cdot\varepsilon_n \longrightarrow 0$, as $n \to \infty$

or $n\cdot\varepsilon_n \to 0$, as $n \to \infty$.

Now, using $\lim_{n\to\infty}\left(1 + \frac{a_n}{n}\right)^n \longrightarrow e^a$ if $\lim_{n\to\infty} a_n = a$,

we have

$$\lim_{n\to\infty} M_{Z_n}(t) = \lim_{n\to\infty}\left[1 + \frac{\frac{t^2}{2} + n\varepsilon_n}{n}\right]^n \longrightarrow e^{\frac{t^2}{2}}, \text{ from } \frac{t^2}{2} + n\varepsilon_n \to \frac{t^2}{2}.$$

Thus, we proved that $M_{Z_n}(t) \longrightarrow e^{t^2/2} = M_Z(t)$, i.e. $Z_n = \frac{S_n - n\mu}{\sqrt{n}\sigma} \longrightarrow N(0,1)$ in distribution.

Simulation study

Let $X_1, X_2, \cdots X_n \sim exp(\lambda)$, $n=5$, $\lambda=2$. $\mu = E(X_1) = \frac{1}{\lambda} = 0.5$

$\sigma^2 = Var(X_1) = \frac{1}{\lambda^2} = 1/4$

distribution of $Z_n = \frac{S_n - n \cdot \mu}{\sqrt{n\sigma^2}} = \frac{S_n - n \cdot \frac{1}{2}}{\sqrt{n \cdot \frac{1}{4}}}$

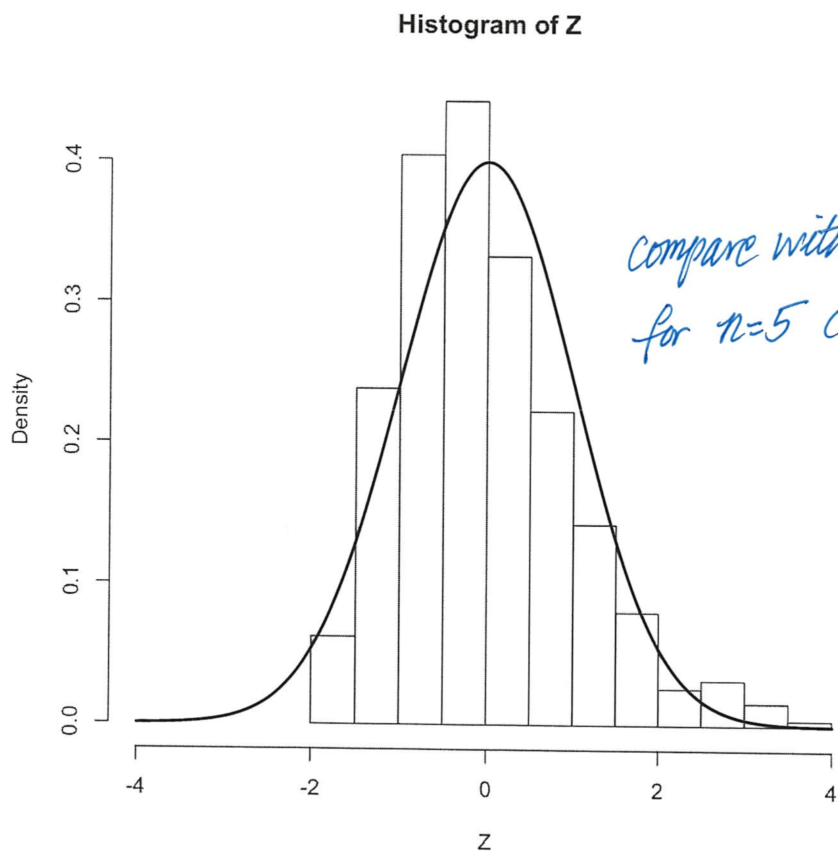
```
> n<-5
> K<-1000
> Z<-rep(0, K)
> for(i in 1:K){
+ X<-rexp(n, 2)
+ Z[i]<-(sum(X)-n*(1/2))/sqrt(n/4)}
> hist(Z, prob=T, xlim=c(-4, 4))
> t<-seq(-4, 4, 0.01)
> lines(t, dnorm(t, 0, 1), lwd=2)
```

generate $X_1, X_2, X_3, X_4, X_5 \sim exp(2)$

$Z[i] = \frac{\sum_{i=1}^{5} X_i - 5 \cdot \frac{1}{2}}{\sqrt{n \cdot \frac{1}{4}}}$, repeat this process $K = 1000$ times.

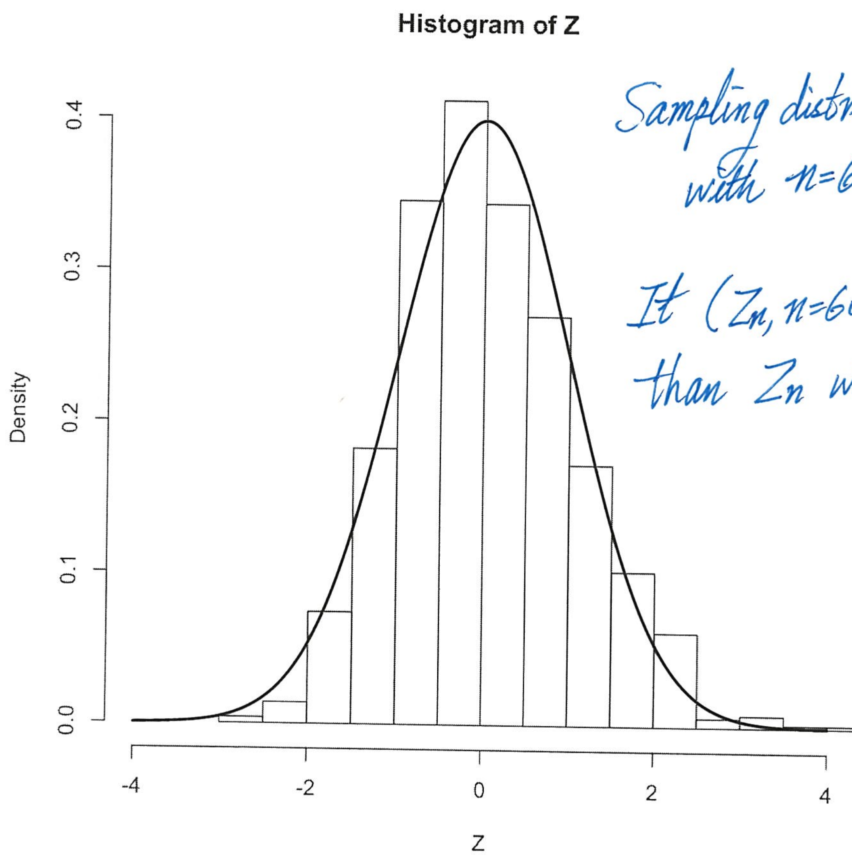
one obtains $Z_1, Z_2, \cdots, Z_{1000}$.

sampling distribution of Z_n based on 1000 sample values.



compare with $N(0,1) = Z$ distribution for $n=5$ case.

```
> n<-60
> K<-1000
> Z<-rep(0, K)
> for(i in 1:K){
+ X<-rexp(n, 2)
+ Z[i]<-(sum(X)-n*(1/2))/sqrt(n/4)}
> hist(Z,  prob=T, xlim=c(-4, 4))
> t<-seq(-4, 4, 0.01)
> lines(t, dnorm(t, 0, 1), lwd=2)
```



Sampling distribution of $Z_n = \frac{S_n - n\mu}{\sqrt{n\sigma^2}}$ with $n=60$.

It (Z_n, $n=60$) is closer to $N(0,1)$ than Z_n with $n=5$.

Exampl D. $X_1, X_2, \cdots, X_n \sim exp(1)$, i.e. $X \sim f(x) = \begin{cases} e^{-x}, & x>0 \\ 0, & x \le 0. \end{cases}$

compare with $X \sim \Gamma(\alpha, \lambda)$ with pdf $f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}$, $x>0$.

one sees $exp(1) = \Gamma(1,1)$. From $\Gamma(\alpha,\lambda) + \Gamma(\beta,\lambda) \sim \Gamma(\alpha+\beta, \lambda)$, one has

$$S_n = X_1 + X_2 + \cdots + X_n = \Gamma(n, 1). \quad (*)$$

From $X \sim \Gamma(\alpha, \lambda)$, $E(X) = \frac{\alpha}{\lambda}$, $Var(X) = \frac{\alpha}{\lambda^2}$, one obtains

$$E(S_n) = \frac{n}{1} = n, \quad Var(S_n) = \frac{n}{1^2} = n.$$

Therefore, from C.L.T. $S_n = \sum_{i=1}^{n} X_i \,\dot{\sim}\, N(E(S_n), Var(S_n))$,

i.e. $S_n \,\dot{\sim}\, N(n, n)$ or $\frac{S_n - n}{\sqrt{n}} \,\dot{\sim}\, N(0,1)$.

From $(*)$, one also has $S_n \sim \Gamma(n,1)$, which is the exact distribution.

Next, we plot CDF (cumulative distribution function) for $N(0,1)$ and standardized gamma distribution for $n=5, 10$, and 30.

For gamma $\Gamma(n,1)$, standardized distribution is

$$F_n(t) = P\left(\frac{S_n - n}{\sqrt{n}} \le t\right) = P(S_n \le n + \sqrt{n}t)$$

$$= P(\Gamma(n,1) \le n + \sqrt{n}t) = pgamma(n+\sqrt{n}t, n, 1).$$

For standard normal dist, $\Phi(t) = P(Z \le t) = pnorm(t, 0, 1)$.

Example D.

```
t<-seq(-4, 4, 0.01)
Phit<-pnorm(t, 0,1)
plot(t, Phit, type="l")
n1<-5
n2<-10
n3<-30
Fn1<-pgamma(n1+sqrt(n1)*t, n1,1)
lines(t, Fn1, lty=1, col="red")
Fn2<-pgamma(n2+sqrt(n2)*t, n2,1)
lines(t, Fn2, lty=2, col="blue")
Fn3<-pgamma(n3+sqrt(n3)*t, n3,1)
lines(t, Fn3, lty=3, col="green")
```

$\leftarrow P(Z \le t) = \Phi(t)$.

$\leftarrow F_{n_1}(t) = P(\Gamma(n_1,1) \le n_1 + \sqrt{n_1}\, t)$

$= \text{pgamma}(n_1 + \sqrt{n_1}\, t, n_1, 1)$

$\leftarrow F_{n_2}(t)$, $n_2 = 10$

$\leftarrow F_{n_3}(t)$, $n_3 = 30$.