Chapter 5. part 3.

More examples

4. $N \sim P(\lambda)$, $E(N)=100$, find Δ such that

$$P(100-\Delta < N < 100+\Delta) \approx 0.9.$$

Ans. $\lambda = E(N) = 100$, notice $N = X_1 + X_2 + \cdots + X_{100}$, each $X_i \sim P(1)$, X_i independent

Thus, $N \,\dot{\sim}\, N(\mu, \sigma^2)$, where $\mu = E(N) = 100$,

$$\sigma^2 = Var(N) = 100 \cdot 1 = 100.$$

From CLT, $N \,\dot{\sim}\, N(100, 10^2)$.

now, $P(100-\Delta < N < 100+\Delta) = P(-\Delta < N-100 < \Delta)$

$$= P\left(-\frac{\Delta}{10} < \frac{N-100}{10} < \frac{\Delta}{10}\right) \approx P\left(-\frac{\Delta}{10} < Z < \frac{\Delta}{10}\right)$$

$$= P\left(|Z| < \frac{\Delta}{10}\right). \qquad \left(\text{from } \frac{N-100}{10} \,\dot{\sim}\, Z\right)$$

if $P(100-\Delta < N < 100+\Delta) \approx 0.9$

i.e. $P\left(|Z| < \frac{\Delta}{10}\right) \approx 0.9$, one has $\frac{\Delta}{10} = 1.645$,

i.e. $\Delta = 16.45$.

Exampl E. Measurement Error.

$X_1, X_2, \cdots, X_n$ iid. $E(X_i) = \mu$, $Var(X_i) = \sigma^2$.

$\bar{X}_n = \frac{1}{n}(X_1 + X_2 + \cdots + X_n)$ is used to estimate μ. $\quad E(\bar{X}_n) = \mu$, $Var(\bar{X}_n) = \sigma^2/n$

C.L.T. $P(|\bar{X}_n - \mu| < c) = P(-c < \bar{X}_n - \mu < c) = P\left(\frac{-c}{\sigma/\sqrt{n}} < \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} < \frac{c}{\sigma/\sqrt{n}}\right)$

$$\approx P\left(-\frac{c}{\sigma/\sqrt{n}} < Z < \frac{c}{\sigma/\sqrt{n}}\right) \qquad \left(\text{b/c } \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}} = \frac{\bar{X}_n - E(\bar{X}_n)}{\sqrt{Var(\bar{X}_n)}} \approx Z = N(0,1)\right)$$

$= \Phi\left(\frac{c}{\sigma/\sqrt{n}}\right) - \Phi\left(-\frac{c}{\sigma/\sqrt{n}}\right)$. eg. $n=16$, $\sigma=1$,

$$P(|\bar{X}_n - \mu| < 0.5) \approx \Phi\left(\frac{0.5}{1/\sqrt{16}}\right) - \Phi\left(-\frac{0.5}{1/\sqrt{16}}\right) = \Phi(2) - \Phi(-2) = 0.954 = 95.4\%.$$

9. Compare binomial cdf and the normal approximation for
(a) $n=20$, $p=0.2$ (b) $n=40$, $p=0.5$.

Ans: In class, one uses R to do the calculations.

binomial $X \sim B(20, 0.2)$.

c.d.f. $F(k) = P(X \le k) = \sum_{m=0}^{k} P(X=m)$

$$= \sum_{m=0}^{k} \binom{20}{m} (0.2)^m (1-0.2)^{20-m}, \quad k=0,1,2,\cdots,20$$

in R, $F(k) = \text{pbinom}(k, 20, 0.2)$.

For normal approximation. $X \dot\sim N(\mu, \sigma^2)$.

$\mu = EX = n \cdot p = 20 \cdot (0.2) = 4$.

$\sigma^2 = Var(X) = n \cdot p \cdot (1-p) = 20 \cdot (0.2) \cdot (1-0.2) = 3.2$.

thus $X \dot\sim N(4, 3.2)$.

c.d.f. $F(x) = P(X \le x) = \text{pnorm}(x, 4, \sqrt{3.2})$ in R.

10. A six-sided die roll 100 times. Using normal approximation to find the prob. that 6 turns up between 15 and 20 times.
Find the prob. that the sum of face values of 100 trials is less than 300.

Ans. Let X be the number of times a six is observed.

$X \sim B(100, \frac{1}{6})$. $EX = n \cdot p = 100 \cdot \frac{1}{6} = 50/3$.

$Var(X) = n \cdot p(1-p) = 100 \cdot (\frac{1}{6})(\frac{5}{6}) = \frac{125}{9}$.

So normal approximation provides

$X \dot\sim N(\frac{50}{3}, \frac{125}{9})$. (or $= P(\frac{15-50/3}{\sqrt{125}/3} \le Z \le \frac{20-50/3}{\sqrt{125}/3}) = 0.4871$

thus $P(15 \le X \le 20) \approx P(\frac{14.5 - \frac{50}{3}}{\frac{\sqrt{125}}{3}} \le Z \le \frac{20.5 - \frac{50}{3}}{\frac{\sqrt{125}}{3}})$

$= P(\frac{-6.5}{\sqrt{125}} \le Z \le \frac{11.5}{\sqrt{125}}) = 0.5677$. 0.4871 (or without continuity correction)

Let $S_{100} = X_1 + X_2 + \cdots + X_{100}$, X_i = i-th face value.

thus $EX_i = 1\cdot\frac{1}{6} + 2\cdot\frac{1}{6} + \cdots + 6\cdot\frac{1}{6} = 3.5$.

$Var(X_i) = E(X_i^2) - (EX_i)^2$

$E(X_i^2) = 1^2\cdot\frac{1}{6} + 2^2\cdot\frac{1}{6} + \cdots + 6^2\cdot\frac{1}{6} = \frac{91}{6}$.

$Var(X_i) = \frac{91}{6} - (3.5)^2 = \frac{35}{12}$.

Thus $E(S_{100}) = 100\cdot E(X_1) = 100\cdot(3.5) = 350$.

$Var(S_{100}) = 100\cdot Var(X_1) = 100\cdot\frac{35}{12} = \frac{3500}{12}$.

from central limit theorem, $S_{100} \sim N(350, \frac{3500}{12})$.

$$P(S_{100} \leq 300) = P\left(\frac{S_{100} - 350}{\sqrt{\frac{3500}{12}}} < \frac{300-350}{\sqrt{\frac{3500}{12}}}\right) = 0.0017.$$

15. bet \$5 on each of a sequence of 50 independent fair games. Use C.L.T. to approximate the prob. that you will lose more than \$75.00.

Ans. Let S_{50} be the amount of money won. Then

$S_{50} = X_1 + X_2 + \cdots + X_{50}$. X_i = amount of money won in i-th game.

X_i	-5	5
p	$\frac{1}{2}$	$\frac{1}{2}$

, $EX_i = (-5)\frac{1}{2} + (5)\frac{1}{2} = 0$

$Var(X_i) = E(X_i^2) - (EX_i)^2$

$= (-5)^2\cdot\frac{1}{2} + 5^2\cdot\frac{1}{2} = 25$.

i.e. $E(X_1) = 0$. $Var(X_1) = 25$. $E(S_{50}) = 50(0) = 0$.

$Var(S_{50}) = 50 \times 25 = 1250$.

thus $S_{50} \sim N(0, 1250)$

$$P(S_{50} < -75) = P\left(\frac{S_{50}}{\sqrt{1250}} < \frac{-75}{\sqrt{1250}}\right) \doteq P\left(Z < \frac{-3}{\sqrt{2}}\right)$$

$= 0.017$.

12. $X_1, X_2, \dots, X_{100} \overset{iid}{\sim} U[-\frac{1}{2}, \frac{1}{2}]$.

$E(X_1)=0$, $Var(X_1)=1/12$.

$S_{100}=X_1+X_2+\dots+X_{100}$, then $E(S_n)=E(S_{100})=100\cdot E(X_1)=0$.

$Var(S_{100})=100\cdot Var(X_1)=100/12$.

From Central Limit Theorem, $S_{100} \,\dot\sim\, N(0, 100/12)$.

i.e. S_{100} is close to normal distribution with mean 0, variance $100/12$.

Now, $P(S_{100}>1) \doteq P\left(\frac{S_{100}-0}{\sqrt{100/12}} > \frac{1-0}{\sqrt{100/12}}\right)$ (Note, $\sqrt{100/12}=5/\sqrt{3}$)

$$=P\left(\frac{S_{100}}{5/\sqrt{3}} > \frac{1}{5/\sqrt{3}}\right) \approx P(Z>0.3469)=0.3645.$$

$$P(S_{100}>2)=P\left(\frac{S_{100}-0}{5/\sqrt{3}} > \frac{2-0}{5/\sqrt{3}}\right)$$

$$\approx P(Z>0.6938)=0.2442.$$

$$P(S_{100}>5)=P\left(\frac{S_{100}-0}{5/\sqrt{3}} > \frac{5-0}{5/\sqrt{3}}\right)$$

$$\doteq P(Z>1.732)\doteq 0.0416.$$

or (prefer?) $P(|S_{100}|>1)=P\left(\frac{|S_{100}|}{5/\sqrt{3}} > \frac{1}{5/\sqrt{3}}\right) \approx P(|Z|>0.3469)$

$$=2*0.3645=0.729.$$

$$P(|S_{100}|>2)=2*0.2442=0.4882.$$

$$P(|S_{100}|>5)=2*0.0416=0.0832.$$

16. $X_1, X_2, \cdots, X_{20} \sim f(x) = 2x,\ 0 \le x \le 1.$

Let $S_{20} = X_1 + X_2 + \cdots + X_{20}$. Use C.L.T to approximate $P(S_{20} \le 10)$.

Ans.

$$E(X_1) = \int_0^1 x \cdot f(x)\,dx = \int_0^1 2x^2\,dx = \frac{2}{3}.$$

$$E(X_1^2) = \int_0^1 x^2 \cdot 2x\,dx = \frac{2}{4} = \frac{1}{2}.$$

$$Var(X_1) = E(X_1^2) - (EX_1)^2 = \frac{1}{2} - \left(\frac{2}{3}\right)^2 = \frac{1}{2} - \frac{4}{9} = \frac{9-8}{18} = \frac{1}{18}.$$

Thus, $E(S_{20}) = 20 \cdot E(X_1) = 20 \cdot \frac{2}{3} = \frac{40}{3}.$

$$Var(S_{20}) = 20 \cdot Var(X_1) = 20 \cdot \frac{1}{18} = \frac{10}{9}.$$

from C.L.T. $S_{20} \,\dot\sim\, N\left(\frac{40}{3}, \frac{10}{9}\right).$

thus, $$P(S_{20} \le 10) = P\left(\frac{S_{20} - \frac{40}{3}}{\sqrt{\frac{10}{9}}} \le \frac{10 - \frac{40}{3}}{\sqrt{\frac{10}{9}}}\right)$$

$$\approx P\left(Z \le \frac{\frac{-10}{3}}{\frac{\sqrt{10}}{3}}\right) = P(Z \le -\sqrt{10})$$

$$= 0.000783.$$

17. X is a measurement, $E(X) = \mu$, $Var(X) = \sigma^2 = 25$.

$\bar{X} = \frac{X_1 + \cdots + X_n}{n}$. find n such that $P(|\bar{X} - \mu| < 1) = 0.95$.

Ans. $\bar{X} \,\dot\sim\, N\left(\mu, \frac{\sigma^2}{n}\right)$ from C.L.T.

thus $\frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$, i.e. $\frac{\bar{X} - \mu}{\frac{5}{\sqrt{n}}} \sim N(0,1)$.

therefore, $$P(|\bar{X} - \mu| < 1) = P\left(\left|\frac{\bar{X} - \mu}{\frac{5}{\sqrt{n}}}\right| < \frac{1}{\frac{5}{\sqrt{n}}}\right)$$

$$= P\left(|Z| < \frac{\sqrt{n}}{5}\right),$$ let $P\left(|Z| < \frac{\sqrt{n}}{5}\right) = 0.95$,

one has $\frac{\sqrt{n}}{5} = 1.96$, $\sqrt{n} = 5 \times 1.96 = 9.8$, $n = 96.04$. n at least 97.

```
###   Question 9 R code
##  (a) n=20, p=0.2. Thus, mu=n*p=4, sig2=n*p*(1-p)=3.2
n<-20
p<-0.2
mu<-n*p
sig<-sqrt(n*p*(1-p))
t<-seq(0, n, 0.01)
plot(t, pbinom(t, n, p), type='l', lwd=2)
lines(t, pnorm(t, mu, sig), lwd=2, col="blue")
title( "Binomial (n, p) vs Normal(mu, sig2)")
##  (b) n=40, p=0.5. Thus, mu=n*p=20, sig2=n*p*(1-p)=10
n<-40
p<-0.5
mu<-n*p
sig<-sqrt(n*p*(1-p))
t<-seq(0, n, 0.01)
plot(t, pbinom(t, n, p), type='l', lwd=2)
lines(t, pnorm(t, mu, sig), lwd=2, col="blue")
title( "Binomial (n, p) vs Normal(mu, sig2)")

###   Question 19 R code
(a) int_0^1cos(2*pi*x)
n<-1000
X<-runif(n, 0,1)
fX<-cos(2*pi*X)
IntF<-mean(fX)
IntF

(b) int_0^1 cos(2*pi*x^2)
n<-1000
X<-runif(n, 0,1)
fX<-cos(2*pi*X^2)          "0.2441"
IntF<-mean(fX)
IntF
```