Lecture 22

11/14/16

- Given joint density function $f(x,y)$, the <u>marginal density of X</u> is $f_X(x) = \int f(x,y)dy$

- (X,Y) independent if $f(x,y) = f_X(x) f_Y(y)$

Thm: If joint density $f(x,y)$ is <u>seperable</u>, ie $f(x,y) = g(x)h(y)$, Then $\exists\, c>0$: $f_X(x) = c\,g(x)$ and $f_Y(y) = \frac{1}{c}h(y)$, and X,Y are independent

5.6 #36 (a)
(b)

Here

- The <u>conditional density of X, given $Y=y$</u> is: $f_X(x|Y=y) = \dfrac{f(x,y)}{f_Y(y)}$ SKIP

5.6 #34

Chap 6: Limit Thms

Sums of discrete RV's: $P(X+Y=z) = \sum_x P(X=x, Y=z-x)$ for any RV X,Y

$= \sum_x P(X=x)P(Y=z-x)$ for independent X,Y

Thm: Let $X \sim Bin(n,p)$ and $Y \sim Bin(m,p)$ with X and Y independent. Then $X+Y \sim Bin(n+m,p)$

Pf: 1. $\leadsto$ $X+Y$ is RV for the # of successes in a sequence of $n+m$ (independent) trials, when each trial has prob. p of success

2. calculate using above formula, and identity $\sum_{j=0}^{k} C_{n,j}\, C_{m,k-j} = C_{n+m,k}$

Similarly,

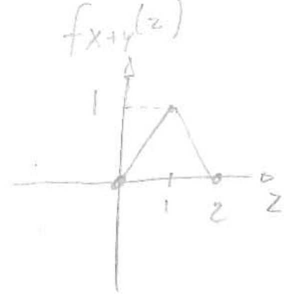
Thm: Let $X \sim Pois(\lambda)$ and $Y \sim Pois(\mu)$ with X and Y independent. Then $X+Y \sim Pois(\lambda+\mu)$

(Pf: Approximate $Pois(\lambda)$ by $Bin(\lfloor n\lambda \rfloor, \frac{1}{n})$ and use previous Thm)

Sums of <u>continuous RV's</u>: $f_{X+Y}(z) = \int f(x, z-x)\,dx$ for any RV X, Y

$$= \int f_X(x) \cdot f_Y(z-x)\,dx \quad \text{for independent RV } X, Y.$$

<u>Ex</u>: X, Y independent and both Uniform on $[0,1]$

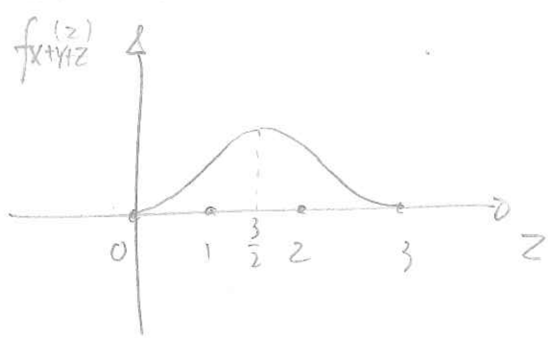


$$f_{X+Y}(z) = \begin{cases} \int_0^z 1 \cdot 1\,dx = z, & \text{if } 0 \le z \le 1 \\ \int_{z-1}^{1} 1 \cdot 1\,dx = 1-(z-1) = 2-z, & \text{if } 1 \le z \le 2 \end{cases}$$

X, Y, Z independent and uniform on $[0,1]$. Let $f_{X+Y}(y) = f_2(y)$

Then $f_{X+Y+Z}(z) = \int f_2(x) \cdot \underbrace{f_Z(z-x)}_{=1 \text{ for } x \text{ ranging between } z-1 \text{ and } z} dx$

$$= \int_{z-1}^{z} f_2(x)\,dx = \begin{cases} \int_0^z x\,dx, & \text{when } 0 \le z \le 1 \to \frac{z^2}{2} \\ \int_{z-1}^{2} (2-x)\,dx, & \text{when } 2 \le z \le 3 \to \frac{(3-z)^2}{2} \\ \int_{z-1}^{1} x\,dx + \int_{z-1}^{2} (2-x)\,dx, & \text{when } 1 \le z \le 2 \overset{\cdots}{\to} z(3-z) - \frac{3}{2} \end{cases}$$



Note: smoothing effect when going $f_X(z) \to f_{X+Y}(z) \to f_{X+Y+Z}(z)$

For now, SKIP Gamma Distribution

Thm: For any RV $X_1, \dots, X_n$: $E(X_1 + \dots + X_n) = E(X_1) + \dots + E(X_n)$

Pf: $n = 2$, discrete RV

$$E(X_1+X_2) = \sum_{x_1,x_2} (x_1+x_2) P(X_1=x_1, X_2=x_2)$$

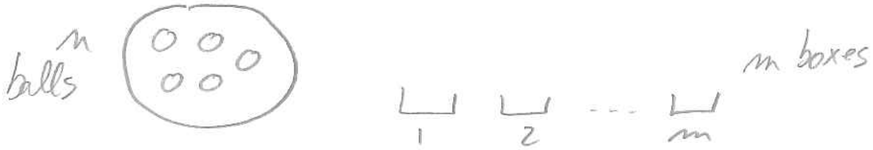
$$= \sum_{x_1,x_2} x_1 P(X_1=x_1, X_2=x_2) + \sum_{x_1,x_2} x_2 P(X_1=x_1, X_2=x_2)$$

Marginals!

$$= \sum_{x_1} x_1 P(X_1=x_1) + \sum_{x_2} x_2 P(X_2=x_2)$$

$$= E(X_1) + E(X_2)$$

Ex: $\boxed{n \text{ balls in } m \text{ boxes}}$ Each ball put in a uniformly chosen box. $E(X)$, where $X = \#$ of empty boxes



Note $X = X_1 + X_2 + \dots + X_m$, where $X_i = \begin{cases} 1 & \text{if box } i \text{ is empty} \\ 0 & \text{otherwise} \end{cases}$

$\xrightarrow{\text{Thm}} E(X) = E(X_1) + \dots + E(X_m)$

Need $E(X_i)$, for all i. These are same for all i, so $E(X) = m\,E(X_1)$

$$E(X_1) = 1 \cdot P(X_1 = 1) + 0 \cdot P(X_1 = 0) = \left(1 - \frac{1}{m}\right)^n \Rightarrow E(X) = m\left(1-\frac{1}{m}\right)^n$$

But, $P(X_1 = 1) = \left(1 - \frac{1}{m}\right)^n$

Do 6.7 #2 Draw 13 cards from 52. Expected # of aces?

Do coupon collector first

Thm: If X and Y independent RV, then $E(XY) = E(X)E(Y)$

Pf $E(XY) = \sum_{x,y} xy\, P(X=x, Y=y) \overset{\text{indep}}{=} \sum_{x,y} xy\, P(X=x) P(Y=y) = \sum_x x P(X=x) \sum_y y P(Y=y) = E(X)E(Y)$

Def: $Cov(X,Y) = E\big((X-E(X))(Y-E(Y))\big) = E(XY) - E(X)E(Y) = \begin{cases} 0 & \text{if } X \text{ and } Y \text{ are indep.} \\ Var(X) & \text{if } Y = X \end{cases}$

Remark: [Claim, that when 13 cards are drawn from deck of 52, the probability that card i is an ace is equal to $\frac{4}{52} = \frac{1}{13}$) for all $i = 1 \ldots 13$

on 6.7#2

Denote $X_i = \begin{cases} 1 & \text{if card } i \text{ is an ace} \\ 0, & \text{otherwise} \end{cases}$. Then $P(\text{card } i \text{ is an ace}) = P(X_i = 1) = E(X_i)$

- $E(X_1) = P(X_1 = 1) = \frac{4}{52} = \frac{1}{13}$
- $E(X_2) = P(X_2 = 1) = \underbrace{P(X_2 = 1 | X_1 = 0)}_{\frac{4}{51}} \overbrace{P(X_1 = 0)}^{\frac{48}{52}} + \underbrace{P(X_2 = 1 | X_1 = 1)}_{\frac{3}{52}} \overbrace{P(X_1 = 1)}^{\frac{4}{52}}$

$$= \frac{4 \cdot 51}{52 \cdot 51} = \frac{1}{13}$$

- $E(X_3) = P(X_3 = 1) = P(X_3 = 1 \mid X_1 = 0, X_2 = 0) P(X_1 = 0, X_2 = 0) + P(X_3 = 1 \mid X_1 = 0, X_2 = 1) P(X_1 = 1, X_2 = 0)$
$+ P(X_3 = 1 \mid X_1 = 1, X_2 = 1) P(X_1 = 1, X_2 = 1)$

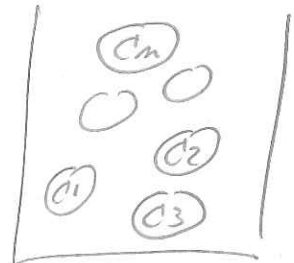
$$= \frac{4}{50} \cdot \frac{48}{52} \cdot \frac{47}{51} + \frac{3}{50} \cdot \left(\frac{48}{52} \cdot \frac{4}{51} + \frac{4}{52} \cdot \frac{48}{51} \right) + \frac{2}{50} \cdot \frac{4}{52} \cdot \frac{3}{51}$$

$$= \frac{4(48 \cdot 47 + 48 \cdot 6 + 6)}{52 \cdot 51 \cdot 50} = \frac{4(48 \cdot 47 + 49 \cdot 6)}{52 \cdot 51 \cdot 50} = \frac{4}{52} \cdot \frac{\cancel{2550}}{\cancel{2550}} = \frac{1}{13}$$

Remark: If we would be picking cards from this deck with replacement, then this would be the expected result (ie that $P(\text{card } i \text{ is an ace}) = \frac{13}{52}$)
What is remarkable is that this is the same probability when there is no replacement!

Coupon Collector's Problem

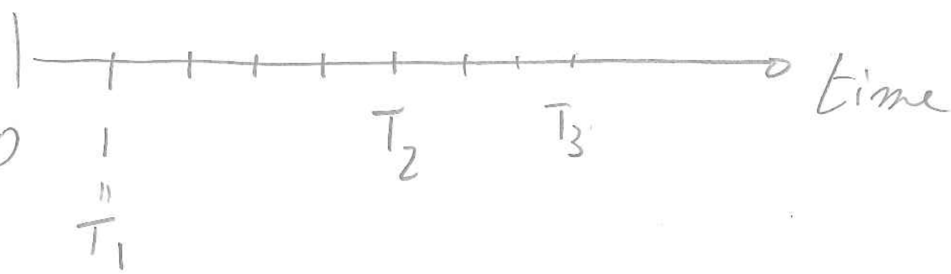
- Urn containing n different coupons
- Each time, a coupon is drawn uniformly from urn; its number is recorded, and then it is placed back in the urn



- Q: <u>Expected # of draws</u> needed to collect all coupons?

$= \text{Expected Time } T$

Let T_i = time needed to collect i^{th} distinct coupon.
Of course $T_1 = 1$



$$T = T_n = \underbrace{(T_n - T_{n-1})}_{X_n} + \underbrace{(T_{n-1} - T_{n-2})}_{X_{n-1}} + \cdots + \underbrace{(T_2 - T_1)}_{X_2} + \underbrace{T_1}_{X_1}$$

<u>Key</u>: Each X_i is a geometric RV with prob of success $p_i = \frac{n-(i-1)}{n}$

ie the fraction of coupons that have not yet been drawn

$$\Rightarrow E(X_i) = \frac{1}{p_i}$$

Thus: $E(T) = E(X_1) + E(X_2) + \cdots + E(X_n)$

$$= \frac{1}{p_1} + \frac{1}{p_2} + \cdots + \frac{1}{p_n}$$

$$= \frac{n}{n} + \frac{n}{n-1} + \frac{n}{n-2} + \cdots + \frac{n}{1} = n\left(\frac{1}{n} + \frac{1}{n-1} + \cdots + \frac{1}{2} + \frac{1}{1}\right) = n\sum_{i=1}^{n} \frac{1}{i} \approx n \ln n$$

as $n \to \infty$

Return to: <u>Thm</u>: $E(XY) = E(X)E(Y)$ if X and Y are independent RV.

<u>Remark</u>: Converse not true!

<u>Ex 6.10</u>:

$Y \backslash X$	0	1
-1	$\frac{1}{4}$	0
0	$\frac{1}{4}$	$\frac{1}{4}$
1	$\frac{1}{4}$	0

$0 = P(X=1, Y=-1) \neq P(X=1)P(Y=-1) = \frac{1}{4} \cdot \frac{1}{4}$

Here

$$Cov(X,Y) \stackrel{def}{=} E\Big((X-E(X))(Y-E(Y))\Big) = E(XY) - E(X)E(Y)$$

<u>Thus</u>: $Cov(X,Y) = 0$ if X and Y are independent

<u>Remark</u>: Converse is not true by Ex above.

<u>Variance of a sum</u>

$$Var(X+Y) = E\Big(\big((X+Y) - E(X+Y)\big)^2\Big) = E\Big(\big[(X-E(X)) + (Y-E(Y))\big]^2\Big)$$

$$= E\big((X-E(X))^2\big) + 2E\Big((X-E(X))(Y-E(Y))\Big) + E\big((Y-E(Y))^2\big)$$

$$\Rightarrow \boxed{Var(X+Y) = Var(X) + 2\,Cov(X,Y) + Var(Y)}$$

More generally: For <u>any</u> RV $X_1, \dots X_n$:

$$Var(X_1 + X_2 + \dots + X_n) = Var(X_1) + Var(X_2) + \dots + Var(X_n)$$
$$+ 2\Big\{Cov(X_1,X_2) + Cov(X_1,X_3) + \dots + Cov(X_1,X_n)$$
$$+ Cov(X_2,X_3) + \dots + Cov(X_2,X_n)$$
$$\dots$$
$$+ Cov(X_{n-1},X_n)\Big]$$

<u>Cor</u>: If $X_1, X_2, \dots, X_n$ are independent, then $Var(X_1 + \dots + X_n) = Var(X_1) + \dots + Var(X_n)$

Lecture 25 11/21/16

Ex ⊛ Bernoulli RV $X = \begin{cases} 1 ; & P(X=1)=p \\ 0 ; & P(X=0)=1-p \end{cases}$ for some $0<p<1$

$E(X) = 1\cdot p + 0\cdot(1-p) = p$

$E(X^2) = 1^2\cdot p + 0^2\cdot(1-p) = p$, so $Var(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1-p)$

⊛ $Y = Bin(n,p)$. Thus $Y = X_1 + X_2 + \dots + X_n$, where each X_i is Bernoulli RV

$E(Y) = n E(X_1) = np$

$Var(Y) \underset{\substack{\uparrow \\ X_i \text{ are independent}}}{=} n Var(X_1) = n p(1-p)$

Do Ex 6.7 #12

One more definition: Standard Deviation of a RV $X = (Var(X))^{1/2}$

Notation: σ, ie $Var(X) = \sigma^2$

Chebyshev's Inequality: For any RV X, if $z>0$:

$$P(|X-E(X)| \geq z) \leq \frac{Var(X)}{z^2}$$

pf: $Var(X) = E((X-E(X))^2) = \int (x-E(X))^2 f(x)\,dx = \int_{|x-E(X)|<z} (x-E(X))^2 f(x)\,dx + \int_{|x-E(X)|\geq z} (x-E(X))^2 f(x)\,dx$

$\geq \int_{|x-E(X)|\geq z} (x-E(X))^2 f(x)\,dx$

$\geq \int_{|x-E(X)|\geq z} z^2 f(x)\,dx$

$= z^2 \int_{|x-E(X)|\geq z} f(x)\,dx$

$= z^2 P(|X-E(X)|\geq z)$

Application: For $z = k\sigma$, where k is a given positive integer, we get

$P(|X-E(X)| \geq k\sigma) \leq \frac{\sigma^2}{k^2\sigma^2} = \frac{1}{k^2}$

ie, the prob. that X deviates more than k standard deviations from $E(X)$, shrinks at a rate of $\frac{1}{k^2}$

$\rightsquigarrow$ motivates the terminology of "standard deviation" for the quantity $\sigma = (Var(X))^{1/2}$

Do Ex 6.7 #18

Law of large numbers

If $X_1, X_2, X_3, \ldots$ is a seq. of i.i.d. RV's with $E(X_i) = \mu$, $Var(X_i) = \sigma^2$,

then the sample mean $\bar{X}_n = \frac{X_1 + X_2 + \cdots + X_n}{n}$ will be close to μ as $n \to \infty$, with high probability!

$$E(\bar{X}_n) = \frac{1}{n} n E(X_1) = \mu$$

$$Var(\bar{X}_n) = \frac{1}{n^2} n E(X_1) = \frac{\sigma^2}{n}$$

Consider: $P(|\bar{X}_n - \mu| \geq \varepsilon)$ for arbitrary $\varepsilon > 0$. By Chebyshev's Inequality:

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

In words: The probability that the sample mean $\bar{X}_n$ deviates at least ε from μ, tends to 0 as $n \to \infty$.

Also, $P(|\bar{X}_n - \mu| < \varepsilon) \to 1$ as $n \to \infty$

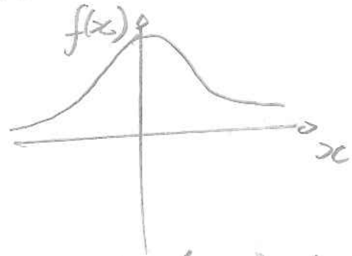
(Strong Law of large numbers: If $X_1, X_2, \ldots$ is a seq of i.i.d. RV's with $E(|X_i|) < \infty$, then $\bar{X}_n \to E(X_1)$ as $n \to \infty$, with probability 1.)

SKIP

Lecture 26 11/23/16

Normal Distribution

Df: Standard normal distribution has density $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, x in $\mathbb{R}$.



Note: symmetry ie $f(-x) = f(x)$

Thm: $f(x)$ is a probability density function.

Pf: Clearly $f(x) \geq 0$. Is $I = \int_{-\infty}^{+\infty} f(x)dx = 1$?

Trick: $I^2 = \int_{-\infty}^{+\infty} f(x)dx \cdot \int_{-\infty}^{+\infty} f(y)dy = \frac{1}{2\pi} \int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} e^{-\frac{x^2+y^2}{2}} dx dy$

$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \end{cases}$ yields $I^2 = \frac{1}{2\pi} \int_0^{\infty}\int_0^{2\pi} e^{-\frac{r^2}{2}} r \, d\theta \, dr$

$\rightarrow I^2 = \frac{2\pi}{2\pi} \int_0^{\infty} e^{-\frac{r^2}{2}} r\,dr = -e^{-\frac{r^2}{2}} \Big|_{r=0}^{r\to\infty} = -0+1$, ie $I^2 = 1$, hence $I = 1$ also.

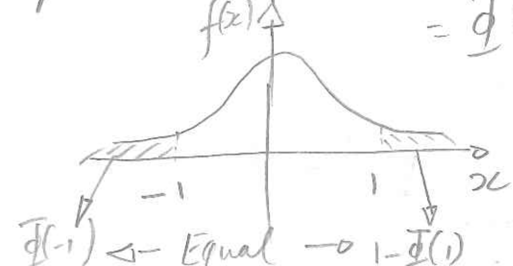
Notation: Distribution function for standard normal RV.

$$\Phi(x) = \int_{-\infty}^{x} f(x)dx, \text{ where } f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

See Table in text on p237

Ex: • $P(2 < X \leq 3) = \Phi(3) - \Phi(2) = 0.9986 - 0.9772 = 0.0214$

• Exploit symmetry: $P(|X| \leq 1) = P(-1 \leq X \leq 1) = \Phi(1) - \Phi(-1)$
$= \Phi(1) - (1 - \Phi(1)) = 2\Phi(1) - 1$



$\boxed{\begin{array}{l} E(X)=0 \\ Var(X)=1 \end{array}}$ • $E(X)=\int_{-\infty}^{+\infty} \underbrace{x f(x)}_{\text{odd function}}\, dx = 0$

• $Var(X)=\int_{-\infty}^{+\infty} \underbrace{x^2 f(x)}_{\text{even function}} dx = \int_0^{+\infty} x^2 f(x) dx$

$= \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} x^2 e^{-\frac{x^2}{2}} dx = -\frac{2}{\sqrt{2\pi}} \int_0^{+\infty} x \, d\left(e^{-\frac{x^2}{2}}\right) = \frac{-2}{\sqrt{2\pi}} x e^{-\frac{x^2}{2}} \Big|_{x=0}^{x\to+\infty} + \frac{2}{\sqrt{2\pi}} \int_0^{+\infty} e^{-\frac{x^2}{2}} dx$

integrate by parts

$= 0 + 2 \cdot \frac{1}{2} = 1$

Creating a NRV with $\left\{\begin{array}{l} \text{mean } \mu \\ \text{variance } \sigma^2 \text{ (or standard deviation } \sigma>0) \end{array}\right\}$

$\to$ Notation: Normal(μ, σ^2)

Set $Y = \sigma X + \mu$, where X is standard normal RV

Then $\left\{\begin{array}{l} E(Y) = \sigma . E(X) + \mu = \sigma . 0 + \mu = \mu \\ Var(Y) = \sigma^2 Var(X) = \sigma^2 . 1 = \sigma^2 \end{array}\right\}$ as desired

Recall: To find the density function of Y, note $Y = r(X)$ where $r(x) = \sigma x + \mu$ is increasing

so $s(Y) = r^{-1}(X) = \frac{1}{\sigma}(Y-\mu)$

and by a previous thm, the density function of Y is given by:

$$f(s(y))\, s'(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2} \cdot \frac{1}{\sigma}$$

$$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}$$

Thm (no proof): If X is Normal(μ, σ_1^2), and X and Y are independent,
Y is Normal(ν, σ_2^2)

then $X+Y$ is Normal$(\mu+\nu, \sigma_1^2+\sigma_2^2)$

$$\begin{array}{r} {}^{1\,1} \\ 9772 \\ 9772 \\ \hline 9544 \end{array}$$

Lecture 27   11/28/16

- Recall SND: $f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \; -\infty < x < +\infty$

  Distribution: $\Phi(x) = \int_{-\infty}^{x} f(s)\,ds$, tabular form $\begin{cases} E(X) = 0 \\ Var(X) = 1 \end{cases}$

<u>Fact 1</u>:
- Percentage of a SND RV that lies within 1 st. deviation

$$P(|X| \le 1) = \Phi(1) - \Phi(-1) = \Phi(1) - (1 - \Phi(1))$$
$$= 2\Phi(1) - 1 = 2.(0.8413) - 1$$
$$\simeq 0.68, \text{ i.e. more than } \frac{2}{3}$$

- Percentage __________ 2 st. deviations;

$$P(|X| \le 2) = 2\Phi(2) - 1 = 2.(0.9772) - 1$$
$$\simeq 0.95, \text{ i.e. about } 95\%$$

- Define NRV Y with mean μ and st. dev. $\sigma > 0$.

$$f(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{y-\mu}{\sigma}\right)^2}, \; -\infty < y < +\infty$$

<u>Ex</u>: Let Y be NRV with $\mu = 1$ and $\sigma = 2$. What is $P(Y \le \;\;)$?

$$P(Y \le 0) = P(Y - 1 \le -1) = P\Big(\underbrace{\frac{Y-1}{2}}_{\overset{\shortparallel}{X}, \text{ a SNRV}} \le -\frac{1}{2}\Big) = P(X \le -\tfrac{1}{2}) = \Phi(-\tfrac{1}{2})$$

<u>CLT</u>

Let $X_1, X_2, \ldots$ be i.i.d. with $\begin{cases} E(X_i) = \mu \\ Var(X_i) = \sigma^2, \text{ where } 0 < \sigma^2 < \infty \end{cases}$. Let $S_n = X_1 + \ldots + X_n$

Then $P\left(a \le \frac{S_n - n\mu}{\sigma\sqrt{n}} \le b\right) \xrightarrow[\text{as } n \to \infty]{} \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx = \Phi(b) - \Phi(a)$

<u>Remark</u>: In Section 6.5, ignore the <u>histogram correction</u>.

Ex: Flip coin 10000 times. What is probability of getting at least 5271 H?

Apply CLT with X_i a Bernoulli RV $\left(\begin{array}{l} =1 \text{ if H, with prob. } 1/2 \\ =0, \text{ otherwise, prob } 1/2 \end{array}\right)$

$$\to \begin{cases} E(X_i) = 1\cdot\frac{1}{2} + 0\cdot\frac{1}{2} = \frac{1}{2} \\ Var(X_i) = E(X_i^2) - (E(X_i))^2 = 1^2\cdot\frac{1}{2} + 0\cdot\frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4} \end{cases}$$

Thus $S_n = \sum_{i=1}^{n} X_n$ counts the # of H. Now $\begin{cases} n\mu = 10000\frac{1}{2} = 5000 \\ \sqrt{n}\,\sigma = 100\cdot\sqrt{\frac{1}{4}} = 50 \end{cases}$

$$P(S_n \geq 5271) = P(S_n - n\mu \geq 5271 - 5000) = P\left(\frac{S_n - n\mu}{\sqrt{n}\,\sigma} \geq \frac{271}{50}\right)$$

$$\overset{CLT}{\approx} P\left(X \geq \frac{271}{50}\right), \text{ where } X \text{ is a SNRV}$$

$$= 1 - \Phi\left(\frac{271}{50}\right)$$

Remark: Exact distribution of S_{10000} is $Bin(10000, \frac{1}{2})$, hence

$P(S_{10000} \geq 5271)$ can be calculated exactly

General Remark: Several examples and problems ask about:

"What is probability of getting exactly x H's?"

$\to$ Using CLT this translates into $P(X = x)$, which is 0 of course. However, the "correct" answer is nonzero and is due to the histogram correction, which we ignore.

Lecture 28 11/30/16

Did 6.7 # 6.7 #30

Lecture 29 12/2/16

Did 6.7 # 37
38
41
43