

MATH 201 (SUMMER 2023, SESH A2)

LECTURE 8 : 05/25/23

ANURAG SAHAY
OFF HRS: BY APPT (VIA ZOOM)

email: anuragsahay@rochester.edu

{ Zoom ID:
979-4693-6650

LECTURES:
9:00 AM - 11:15 AM (ET)
M, T, W, R

COURSE

WEB PAGE

<https://people.math.rochester.edu/grads/asahay/summer2023/math201/index.html>

ALL PHOTOS TAKEN
FROM TEXTBOOK

ANNOUNCEMENTS

LECTURE 7 ↘

- ① NEW H.W. IS UPLOADED., NO CLASS ON MONDAY (MEMORIAL DAY)
- ② OFFICE HOURS : TODAY - AFTER CLASS. (11:15 - 12:15)
- ③ UPCOMING DEADLINES :
 - Ⓐ WW04 - FRI, MAY 26th
 - Ⓑ HW03 - SAT, MAY 27th
 - Ⓒ WW05 - TUES, MAY 30th
 - Ⓓ HW04 - TUES, MAY 30th } → TO BE UPLOADED
- ④ ERROR IN HW2 - SEE ANNOUNCEMENT ON BLACKBOARD
- ⑤ DETAILS ABOUT EXAMS → SEE WEBPAGE. → IN CLASS
MIDTERM - 2h
FINAL - 3h
- ⑥ PLEASE KEEP VIDEOS ON, IF POSSIBLE !

§ 3.5 NORMAL DISTRIBUTION

SO FAR :

WE HAVE ONLY SEEN ONE EXAMPLE OF
A CONTINUOUS R.V. (i.e. Unif $[a, b]$)

MOST IMP.
CONCEPT



NORMAL DISTRIBUTION.
CONT. R.V.

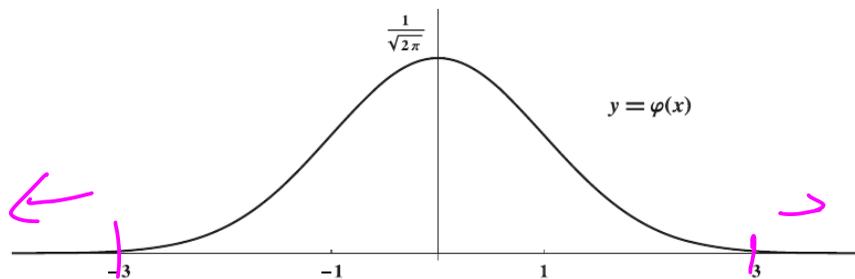
C.L.T. → MANY DISTRIB.
ARE WELL-APPROX.
BY NORMALS.

Definition 3.55. A random variable Z has the **standard normal distribution** (also called **standard Gaussian distribution**) if Z has density function

$$\varphi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \quad (3.39)$$

} p.d.f.

on the real line. Abbreviate this by $Z \sim \mathcal{N}(0, 1)$.



BELL CURVE

Figure 3.9. The probability density function φ of the standard normal distribution.

NOTES :

① A p.d.f SATISFIES

$$\int_{-\infty}^{\infty} f(x) dx = P(X \in \mathbb{R}) = \underline{1}$$

$z \rightarrow \phi$

$$\int_{-\infty}^{\infty} e^{-x^2/2} dx = \sqrt{2\pi}$$

TO BE
DONE NEXT
WEEK

(2) C.D.F? → APPENDIX E OF THE BOOK

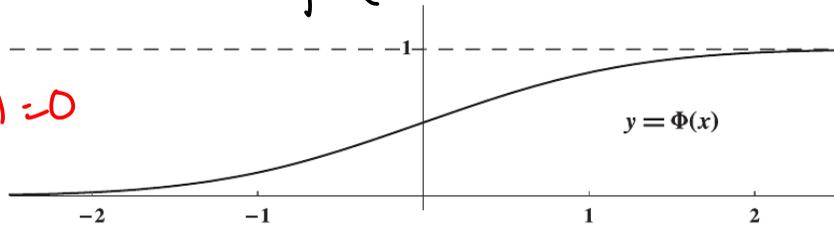
$$\Phi(t) = P(Z \leq t) = \int_{-\infty}^t \frac{1}{\sqrt{2\pi}} e^{-x^2/2} dx$$

↓
erf(t)

$$= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^t e^{-x^2/2} dx$$

ANTI-DERIV. OF $e^{-x^2/2}$

$\lim_{x \rightarrow -\infty} \Phi(x) = 0$



$\lim_{x \rightarrow \infty} \Phi(x) = 1$

Figure 3.10. The cumulative distribution function Φ of the standard normal distribution. It satisfies $0 < \Phi(x) < 1$ for all x .

③

ϕ

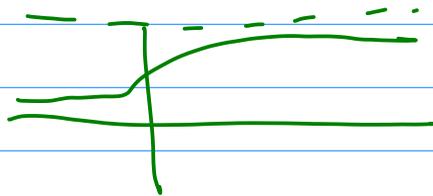
Φ

Φ

} p.d.f & c.d.f.
of $N(0, 1)$

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$\Phi \rightarrow$ APPENDIX E



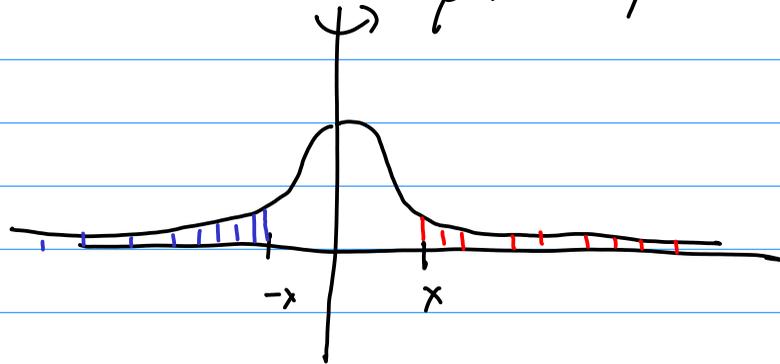
④

$$\Phi(x) + \Phi(-x) = 1$$

$$\phi(x) = \phi(-x)$$

$$\Phi(-x) = \int_{-\infty}^{-x} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$

$$= \int_x^{\infty} \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$



$$\bar{\Phi}(x) = \int_{-A}^x \phi(x) dx$$

$$\bar{\Phi}(x) + \bar{\Phi}(-x) = \int_{-\infty}^x \phi(x) dx + \int_x^{\infty} \phi(x) dx$$

$$= \int_{-\infty}^{\infty} \phi(x) dx = 1$$

p.d.f.

$$\therefore \bar{\Phi}(-x) = 1 - \bar{\Phi}(x)$$

⑤ APPENDIX E

$\Phi(t)$ FOR $0 \leq t \leq 3.49$ [$t \sim 2$ DECIMAL]

2nd DECIMAL OF t

1st DECIMAL OF t

	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817

$$\Phi(1.17) \approx 0.8790$$

$$\Phi(1.2) \approx 0.8849$$

$$\begin{aligned} \Phi(-0.03) &= 1 - \Phi(0.03) \\ &\approx 1 - 0.5120 \\ &= 0.4880 \end{aligned}$$

Example 3.57. Let $Z \sim \mathcal{N}(0, 1)$. Find the numerical value of $P(-1 \leq Z \leq 1.5)$.

$$P(-1 \leq Z \leq 1.5) = P(-1 < Z \leq 1.5)$$

$$P(Z = -1) = 0$$

$$= P(Z \leq \underline{\underline{1.5}}) - P(Z \leq \underline{\underline{1}})$$

$$= \Phi(1.5) - \Phi(1)$$

$$\approx 0.9332 - 0.8413$$

$$= 0.0919$$

$$\Phi(z) \approx \frac{1}{2} \left[\frac{2}{3} + 1 \right]$$

$$= \frac{5}{6} \approx 0.833$$

$$\left(2\Phi(z) - 1 \approx \frac{2}{3} \right)$$

$$\Phi(0.97) \approx 0.8340, \quad \Phi(0.96) \approx 0.8315$$

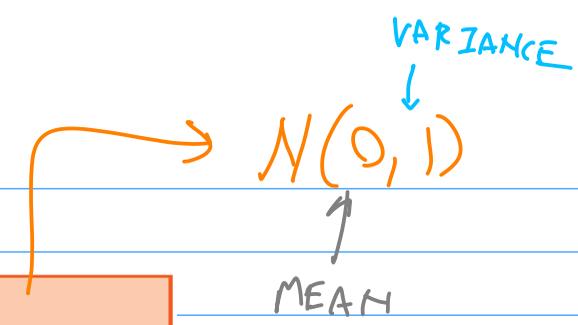
$$0.96 < z < \underline{0.97}$$

$$\boxed{z \approx 0.97}$$

$$(-0.97, 0.97)$$

MEAN & VARIANCE

Fact 3.59. Let $Z \sim \mathcal{N}(0, 1)$. Then $E(Z) = 0$ and $\text{Var}(Z) = E(Z^2) = 1$.

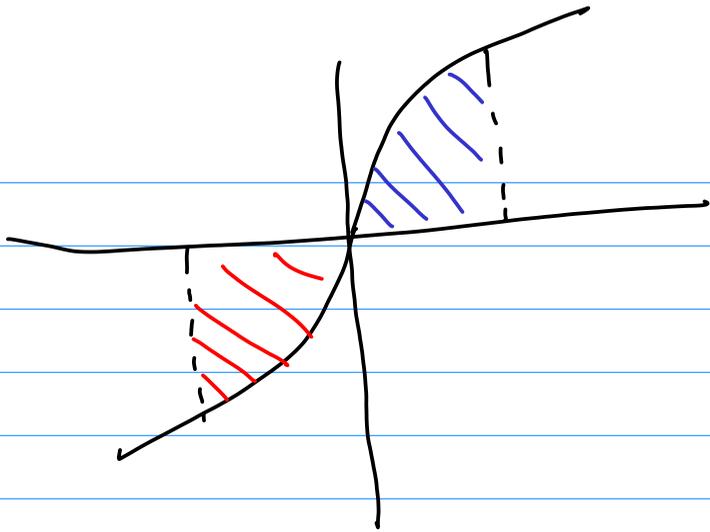


Pf.

$$E(Z) = \int_{-\infty}^{\infty} x \phi(x) dx = \int_{-\infty}^{\infty} x \frac{e^{-x^2/2}}{\sqrt{2\pi}} dx$$

$$= \lim_{a \rightarrow \infty} \int_{-a}^a \frac{x e^{-x^2/2}}{\sqrt{2\pi}} dx = \lim_{a \rightarrow \infty} 0 = 0$$

ODD FUNCTION



$$\text{Var}(z) = E(z^2) - E(z)^2$$

$$= \int_{-\infty}^{\infty} \frac{x^2 e^{-x^2/2}}{\sqrt{2\pi}} dx = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \underbrace{x}_{u} \cdot \underbrace{(x e^{-x^2/2})}_{dv} dx$$

$$\int \overbrace{x}^u \cdot \overbrace{(x e^{-x^2/2}) dx}^{dv} = 1 \cdot (-e^{-x^2/2}) - \int (-e^{-x^2/2}) dx$$

$$du = dx \quad v = -e^{-x^2/2}$$

$$\frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-x^2/2} dx = \left[-e^{-x^2/2} \right]_{-\infty}^{\infty} + \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-x^2/2} dx$$

$$= 0$$

$$\lim_{x \rightarrow \pm\infty} e^{-x^2/2} = 0$$

I.

$$E(z^2) = 1$$

$$\text{Var}(z) = 1 - 0 = 1$$



AFFINE
TRANSFORM

LET

$$X = \sigma Z + \mu$$

$$Z \sim \mathcal{N}(0, 1)$$

$$\sigma > 0, \mu \in \mathbb{R}$$

$E(X)$, $\text{Var}(X)$, f_X , F_X ?

$$E(X) = E(\sigma Z + \mu) = \sigma \cancel{E(Z)} + \mu = \mu$$

$$\text{Var}(X) = \text{Var}(\sigma Z + \mu) = \sigma^2 \boxed{\text{Var}(Z)} = \sigma^2$$

$$F_X(t) = P(X \leq t)$$

$$Z \sim N(0, 1)$$

$$= P(\sigma Z + \mu \leq t)$$

$$= P(\sigma Z \leq t - \mu)$$

$$= P\left(Z \leq \frac{t - \mu}{\sigma}\right)$$

$$F_X(t) = \Phi\left(\frac{t - \mu}{\sigma}\right)$$

$Z \rightarrow \text{CONT.}, X \rightarrow \text{CONT.}$

$$\Rightarrow f_X(t) = F_X'(t)$$

$$= \frac{d}{dt} \Phi\left(\frac{t-\mu}{\sigma}\right)$$

$$= \frac{d}{dt} \left[\frac{t-\mu}{\sigma} \right] \cdot \Phi'\left(\frac{t-\mu}{\sigma}\right)$$

[CHAIN
RULE]

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

FACT .

$$\phi = \Phi'$$

Definition 3.60. Let μ be real and $\sigma > 0$. A random variable X has the *normal distribution with mean μ and variance σ^2* if X has density function

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad (3.44)$$

on the real line. Abbreviate this by $X \sim \mathcal{N}(\mu, \sigma^2)$.

$$E(X) = \mu, \quad \text{Var}(X) = \sigma^2$$

RULE OF THUMB \rightarrow STANDARDIZE.

FACT: $X \sim \mathcal{N}(\mu, \sigma^2)$, THEN $\frac{X - \mu}{\sigma} = Z \sim \mathcal{N}(0, 1)$

Pf: REVERSE $X = \sigma Z + \mu$

Example 3.62. Suppose $X \sim \mathcal{N}(-3, 4)$. Find the probability $P(X \leq -1.7)$.

$\mu = -3, \sigma^2 = 4$ ($\sigma = 2$).

$P(X \leq -1.7)$

$X \leq -1.7 \iff Z \leq \dots$

$Z = \frac{X - \mu}{\sigma} = \frac{X - (-3)}{2} = \frac{X + 3}{2}$

$$X \leq -1.7 \Leftrightarrow X + 3 \leq 3 - 1.7$$

$$\Leftrightarrow \frac{X + 3}{2} \leq \frac{3 - 1.7}{2} = \frac{1.3}{2} = 0.65$$

$$P(X \leq -1.7) = P\left(\frac{X + 3}{2} \leq 0.65\right) = \Phi(0.65) = 0.7422$$

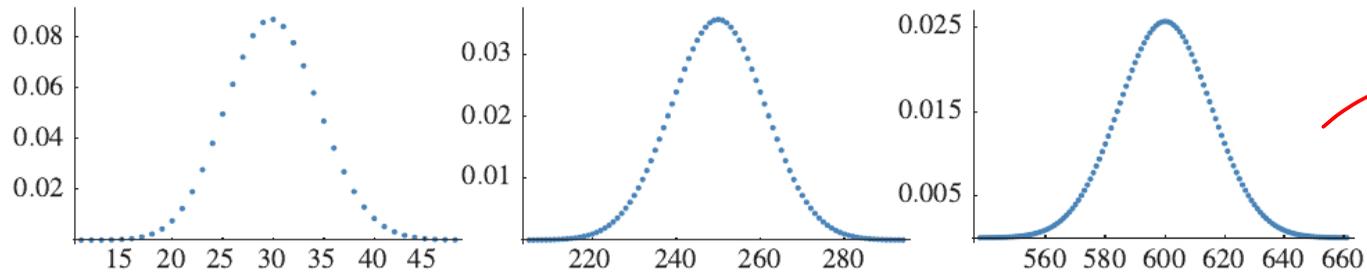
$\sim N(0,1)$

BREAK TILL

10:15 AM ET

§ 4.1 NORMAL APPROXIMATION

Q WHAT MAKES $N(\mu, \sigma^2)$ INTERESTING?



BELL CURVE.

$n=100$
 $p=0.3$



Figure 4.1. The probability mass function of the $\text{Bin}(n, p)$ distribution with parameter values $(100, 0.3)$, $(500, 0.5)$, and $(1000, 0.6)$. In each case the graph shows the function on the interval $(np - 4\sqrt{np(1-p)}, np + 4\sqrt{np(1-p)})$ and the scale on the y-axis changes also.

SEEMS LIKE $\text{Bin}(n, p) \approx N(\mu, \sigma^2)$

AS $n \rightarrow \infty$

$\mu = \text{MEAN OF Bin}(n, p)$, $\sigma^2 = \text{VARIANCE OF Bin}(n, p)$

$$= np$$

$$= npq$$

$$= np(1-p)$$

($q := 1-p$)

$$\text{Bin}(n, p) \approx N(np, np(1-p))$$

$$\begin{aligned} \mu &= n \cdot p \\ &= 1000 \times 0.6 \\ &= 600 \end{aligned}$$

$$\begin{aligned} \sigma^2 &= n \cdot p \cdot (1-p) \\ &= (1000)(0.6)(0.4) \\ &= 240 \end{aligned}$$

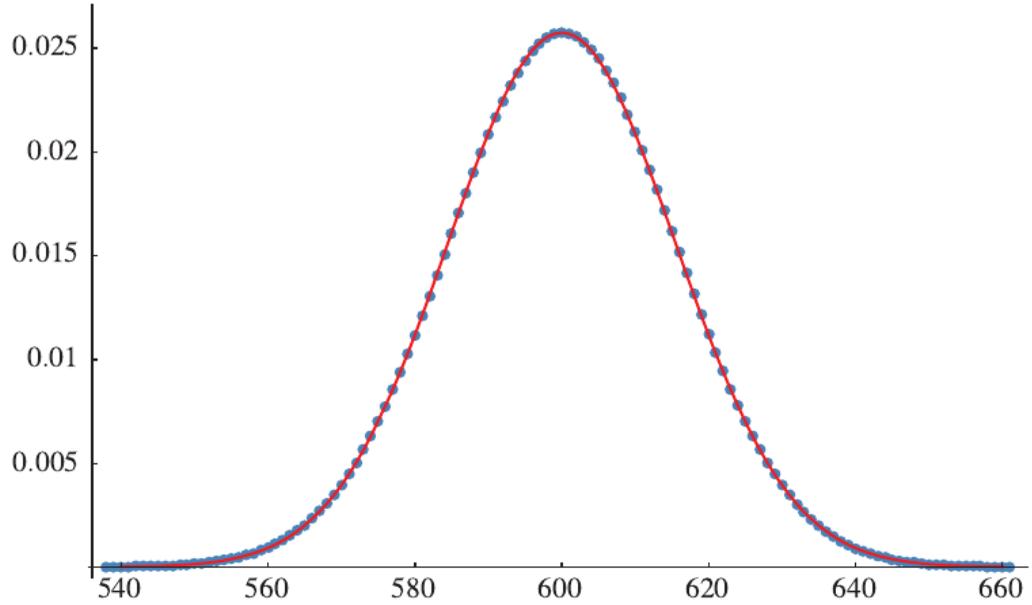


Figure 4.3. The bullets represent the probability mass function of $\text{Bin}(1000, 0.6)$ and the continuous curve is the density function of $\mathcal{N}(600, 240)$ in the interval $[540, 660]$.

$$S_n \sim \text{Bin}(n, p), X \sim \mathcal{N}(np, npq) \Rightarrow S_n \xrightarrow[n \rightarrow \infty]{AS} X$$

PROBLEM :

S_n & X BOTH
HAVE GROWING μ, σ

FIX : RENORMALIZE. / STANDARDIZE.

FOR Y w/ $E(Y) = \mu, \text{Var}(Y) = \sigma^2$

$$\hat{y} = \frac{Y - \mu}{\sigma} \quad \leftarrow \begin{array}{l} \text{AFFINE} \\ \text{OF} \\ Y \end{array} \quad \begin{array}{l} \text{FUNCTION} \\ \text{OF} \\ Y \end{array}$$

$$E(\hat{y}) = \frac{E(Y) - \mu}{\sigma} = \frac{\mu - \mu}{\sigma} = 0 \quad \left| \quad \begin{array}{l} \text{Var}(\hat{y}) = \frac{1}{\sigma^2} \text{Var}(Y) \\ = \frac{1}{\sigma^2} \cdot \sigma^2 = 1 \end{array} \right.$$

$$Y \longrightarrow \hat{Y} = \frac{Y - \mu}{\sigma}$$

$$E(Y) = \mu$$

$$\text{Var}(Y) = \sigma^2$$

$$E(\hat{Y}) = 0$$

$$\text{Var}(\hat{Y}) = 1$$

$$S_n \sim \text{Bin}(n, p) \Rightarrow \hat{S}_n = \frac{S_n - E(S_n)}{\sqrt{\text{Var}(S_n)}} = \frac{S_n - np}{\sqrt{np(1-p)}}$$

CLT FOR Bin $(n, p) \approx \hat{S}_n \longrightarrow N(0, 1)$

Theorem 4.1. (Central limit theorem for binomial random variables)

Let $0 < p < 1$ be fixed and suppose that $S_n \sim \text{Bin}(n, p)$. Then for any fixed $-\infty \leq a \leq b \leq \infty$ we have the limit

$$\lim_{n \rightarrow \infty} P \left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b \right) = \int_a^b \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} dx. \quad (4.1)$$

$$P \left(a \leq \hat{S}_n \leq b \right) \xrightarrow{(n \rightarrow \infty)} P \left(a \leq Z \leq b \right)$$

Pf \rightarrow SEE BOOK.

Normal approximation of the binomial distribution. Suppose that $S_n \sim \text{Bin}(n, p)$ with n large and p not too close to 0 or 1. Then

$$P\left(a \leq \frac{S_n - np}{\sqrt{np(1-p)}} \leq b\right) \text{ is close to } \Phi(b) - \Phi(a). \quad (4.3)$$

As a rule of thumb, the approximation is good if $np(1-p) > 10$.

$$P(a \leq \hat{S}_n \leq b)$$

\parallel

$$\begin{aligned} P(a \leq Z \leq b) \\ = \Phi(b) - \Phi(a) \end{aligned}$$

Example 4.2. A fair coin is flipped 10,000 times. Estimate the probability that the number of heads is between 4850 and 5100.

Example 4.2. A fair coin is flipped 10,000 times. Estimate the probability that the number of heads is between 4850 and 5100.

$S \rightarrow$ # OF HEADS

$$S \sim \text{Bin} \left(\underbrace{10000}_n, \underbrace{\frac{1}{2}}_p \right)$$

$$P(4850 \leq S \leq 5100)$$

$$\mu = E(S) = np$$

$$= (10000) \cdot \frac{1}{2} = 5000$$

$$\sigma^2 = \text{Var}(S) = n \cdot p(1-p) = (10000) \left(\frac{1}{2}\right) \left(\frac{1}{2}\right) = 2500$$

$$S \approx N(5000, 2500)$$

$$\Rightarrow \sigma = \sqrt{2500} = 50$$

$$P(4850 \leq S \leq 5100)$$

$$\frac{4850 - 5000}{50} \leq \frac{S - 5000}{50} \leq \frac{5100 - 5000}{50}$$

$$\begin{array}{c} \parallel \\ -3 \end{array}$$

$$\begin{array}{c} \downarrow \\ 2 \end{array}$$

$$P\left(-3 \leq \frac{S - 5000}{50} \leq 2\right) \approx \Phi(2) - \Phi(-3)$$

↓ c.v. 7.

✓
S

$$\approx 0.9772 - (1 - 0.9987)$$

$$\approx 0.9759$$

$$P(4850 \leq S \leq 5100) = \sum_{4850 \leq k \leq 5100} \binom{10000}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{10000-k}$$

p.m.f of
Bin(10000, 1/2)

$$= \sum_{k=4850}^{5100} \binom{10000}{k} \cdot 2^{-10000}$$

ACTUAL = 0.9765

APPROX = 0.9759

} 0.06%

JK 17

Example 4.3. Suppose a game piece moves along a board according to the following rule: if a roll of a die gives 1 or 2 then take two steps ahead, otherwise take three steps. Approximate the probability that after 120 rolls the piece has moved more than 315 steps.

CONTINUITY CORRECTION

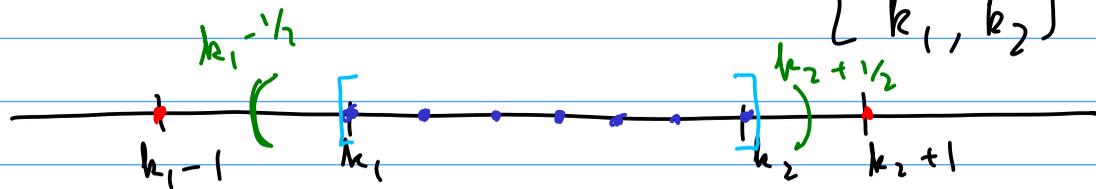
$$S_n \sim \text{Bin}(n, p) \Rightarrow S_n \in \{0, \dots, n\}$$



INTEGER-VALUED.

$$\therefore P(k_1 \leq S_n \leq k_2) = P(k_1 - 1/2 \leq S_n \leq k_2 + 1/2)$$

$(k_1, k_2 \in \mathbb{Z})$



$$[k_1, k_2] \cap \mathbb{Z} = [k_1 - 1/2, k_2 + 1/2] \cap \mathbb{Z}$$

$$P(k_1 \leq S_n \leq k_2) = P(k_1 - 1/2 \leq S_n \leq k_2 + 1/2)$$

$$\approx P(k_1 \leq X \leq k_2)$$

$$X \sim N(np, np(1-p))$$

$$P(k_1 - 1/2 \leq X \leq k_2 + 1/2)$$

BETTER APPROX.

CONT. CORRECTION.

Example 4.5. Roll a fair die 720 times. Estimate the probability that we have exactly 113 sixes.

$X \rightarrow$ # OF SIXES.

$$\mu = 720 \cdot \frac{1}{6} = 120$$

$$X \sim \text{Bin} \left(720, \frac{1}{6} \right)$$

$$\sigma^2 = 720 \cdot \frac{1}{6} \cdot \frac{5}{6} = 100$$

$$P(X = 113) = P(112.5 \leq X \leq 113.5)$$

$$\frac{X - 120}{10}$$

$\nearrow \mu$
 $\downarrow \sigma$

$$= P \left(\underbrace{\frac{112.5 - 120}{10}}_{-0.75} \leq \underbrace{\frac{X - 120}{10}}_Z \leq \underbrace{\frac{113.5 - 120}{10}}_{-0.65} \right)$$

$$\begin{aligned} P(X=113) &= P(-0.75 \leq \hat{X} \leq -0.65) \\ &\approx P(-0.75 \leq Z \leq -0.65) \\ &= \Phi(-0.65) - \Phi(-0.75) \\ &= (1 - 0.7422) - (1 - 0.7734) \\ &= 0.0312 \end{aligned}$$

$P_X(113) = \binom{720}{113} \frac{5^{607}}{6^{720}}$ (p.m.f. of $\text{Bin}(n, p)$)
 ≈ 0.0318 $\left(\binom{n}{k} \cdot p^k \cdot (1-p)^{n-k} \right)$

RED 9

Example 4.2. A fair coin is flipped 10,000 times. Estimate the probability that the number of heads is between 4850 and 5100.

(ACCOUNT
FOR
C.C.)

$$\mu = 10000 \cdot \frac{1}{2} = 5000$$
$$\sigma = \sqrt{10000 \cdot \frac{1}{2} \cdot \frac{1}{2}} = 50$$

$$P(4850 \leq S \leq 5100) = P(4849.5 \leq S \leq 5100.5)$$
$$= P\left(\frac{4849.5 - 5000}{50} \leq \frac{S - 5000}{50} \leq \frac{5100.5 - 5000}{50}\right)$$
$$\approx \Phi(2.01) - \Phi(-3.01)$$

$$\approx \underbrace{\Phi(2.01)}_{0.9778} - \underbrace{\left(1 - \Phi(-3.01)\right)}_{\Phi(3.01)} \\ = 1 - \Phi(3.01)$$

$$= 0.9765$$

ACTUAL = 0.9765

APPROX = 0.9759
(w/o CORR.)

APPROX. = 0.9765
(w/ CORR.)

REMINDER : NO CLASS
ON MONDAY