Chapter 5. Distributions of Functions of Random Variables

Math 3215 Spring 2024

Georgia Institute of Technology

Section 1.
Functions of One Random Variable

Functions of One Random Variable

Let X be a random variable.

Define $Y = u(X)$ for some function u.

We discuss how to find the distribution of Y from that of X.

Functions of One Random Variable

Example

Let X have a discrete uniform distribution on the integers from -2 to 5.

Find the distribution of $Y = X^2$.

$$X = \underbrace{-2, -1, 0, 1, 2, 3, 4, 5}_{8}$$

8 possible values with the same prob.

$$f_X(x) = \frac{1}{8} \quad \text{for} \quad x = -2, -1, \cdots, 5.$$

$$Y = X^2 = (-2)^2, (-1)^2, 0^2, 1^2, 2^2, 3^2, 4^2, 5^2$$

$$f_Y(y) = \begin{cases} \frac{2}{8}, & y = 1, 4 \\ \frac{1}{8}, & y = 0, 9, 16, 25 \end{cases}$$

Recall f is increasing

$\Leftrightarrow$ if $x \leq y$ then $f(x) \leq f(y)$

CDF Technique

Example

Let X have a gamma distribution with PDF

$$f(x) = \frac{1}{\Gamma(\alpha)\theta^\alpha} x^{\alpha-1} e^{-\frac{x}{\theta}}.$$

Find the distribution of $Y = e^X$. > 0

CDF, PDF.

Possible values for Y?

$F_Y(y) = \mathbb{P}(Y \leq y) = 0$ if $y \leq 0$

For $y > 0$,

$F_Y(y) = \mathbb{P}(Y \leq y) = \mathbb{P}(e^X \leq y) = \mathbb{P}(X \cdots)$

$= \mathbb{P}(\ln(e^X) \overset{?}{\leq} \ln y)$, where $\ln(e^X) = X$

Yes? why? $\ln(\cdot)$ is increasing

$= \mathbb{P}(X \leq \ln y) = F_X(\ln y)$

$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy}(F_X(\ln y)) = f_X(\ln y) \cdot (\ln y)'$

Chain Rule

$$= \frac{1}{\Gamma(\theta)\,\theta^{\alpha}} (\ln y)^{\alpha-1} e^{-\frac{1}{\theta}\ln y} \cdot \frac{1}{y}.$$

PDF of Y

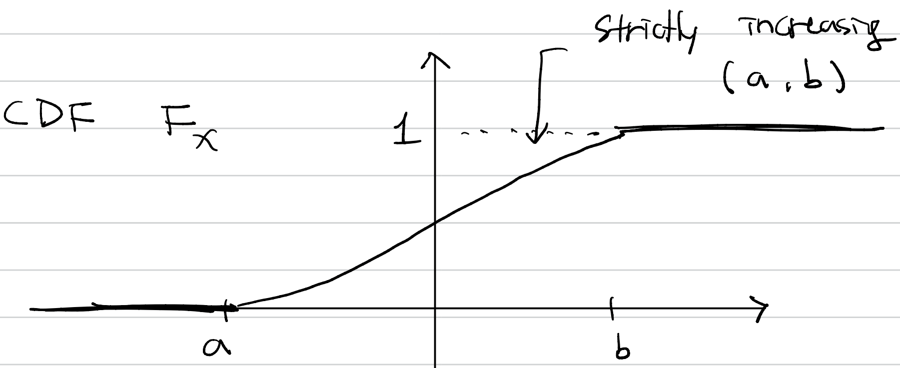
$$e^{-\frac{1}{\theta}\ln y} = y^{-\frac{1}{\theta}}$$

$$y \cdot e^{-\frac{1}{\theta}} = e^{-\frac{1}{\theta} + \ln y}$$

$$e^{\left(-\frac{1}{\theta}\right)\cdot \ln y} = e^{\ln\left(y^{-\frac{1}{\theta}}\right)} = y^{-\frac{1}{\theta}}$$

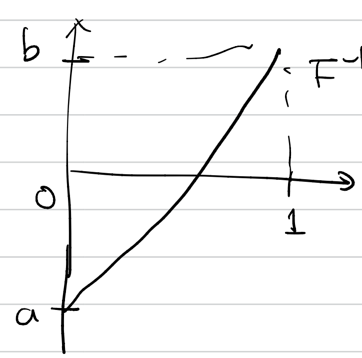
$Y \sim \text{Unif}(0,1)$

X a RV CDF F_X



(f is strictly increasing if
$(x < y \Rightarrow f(x) < f(y))$)

$F^{-1} \leftarrow$ $F(F^{-1}(x)) = x = F^{-1}(F(x))$
$\uparrow$ the inverse of F on (a,b)



$$Z = F_X^{-1}(Y)$$

$$\begin{aligned} F_Z(z) &= P(Z \le z) \\ &= P(F_X^{-1}(Y) \le z) \\ &= P(F_X(F_X^{-1}(Y)) \le F_X(z)) \\ &= P(Y \le F_X(z)) = F_X(z) \end{aligned}$$

F is increasing

X, Z have the same dist.

CDF Technique

Theorem

Let X be a random variable with CDF F.

Suppose F is strictly increasing, $F(a) = 0$, $F(b) = 1$.

Let $Y \sim U(0,1)$.

Then, $X \overset{d}{=} F^{-1}(Y)$.

" $\overset{d}{=}$ " : have the same distribution

<u>Example</u> Q : How to sample random number which follows Exp(2) ?

① Sample x from Unif(0,1)

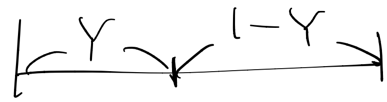
② $F^{-1}(x)$ $\quad$ $F(x)$: CDF of Exp(2)

$\quad = -\frac{1}{2}\ln(1-x)$ $\quad\quad$ $= 1-e^{-2x}$

$$1-(1-X)^3 \sim \text{Unif}(0,1)$$

$$1-Y \sim \text{Unif}(0,1)$$

$$Y \sim \text{Unif}(0,1)$$



Change of Variables

$$F_X(x) = 1-(1-x)^3$$

Example

Let X have the PDF $f(x) = 3(1-x)^2$ for $0 < x < 1$.

Find the distribution of $Y = (1-X)^3$. $Y \in (0,1)$

For $0 < y < 1$,

$$F_Y(y) = P(Y \leq y) = P((1-X)^3 \leq y)$$

$$= P(((1-X)^3)^{\frac{1}{3}} \leq y^{\frac{1}{3}})$$

$$= P(1-X \leq y^{\frac{1}{3}})$$

$$= P(X \geq 1-y^{\frac{1}{3}}) = 1 - F_X(1-y^{\frac{1}{3}})$$

$$f_Y(y) = -f_X(1-y^{\frac{1}{3}}) \cdot \left(-\frac{1}{3} \cdot y^{-\frac{2}{3}}\right)$$

$$= -\cancel{3} \cdot (1-(1-y^{\frac{1}{3}}))^2 \left(-\frac{\cancel{1}}{\cancel{3}} y^{-\frac{2}{3}}\right)$$

$$= (y^{\frac{1}{3}})^2 \cdot y^{-\frac{2}{3}} = 1 \quad \text{for} \quad y \in (0,1)$$

$$Y \sim \text{Unif}(0,1)$$

3 random points uniformly

Indep.

X smallest.

Dist ?

$U_1, U_2, U_3 \sim \text{Unif}(0,1]$ Indep.

$$X = \min\{U_1, U_2, U_3\}$$

$$\mathbb{P}(X \leq t) = 1 - (1-t)^3 = F_X(t)$$

$$\mathbb{P}(X \leq \tfrac{1}{3}) \overset{?}{=} 1 - \mathbb{P}(X > \tfrac{1}{3})$$

$$= 1 - \mathbb{P}(U_1 > \tfrac{1}{3}, U_2 > \tfrac{1}{3}, U_3 > \tfrac{1}{3})$$

$$= 1 - \mathbb{P}(U_1 > \tfrac{1}{3})\,\mathbb{P}(U_2 > \tfrac{1}{3})\,\mathbb{P}(U_3 > \tfrac{1}{3})$$

$$= 1 - (1 - \tfrac{1}{3})^3$$

Exercise

Let X have the PDF $f(x) = 4x^3$ for $0 < x < 1$.

Find the PDF of $Y = X^2$.

Section 2.
Transformations of Two Random Variables

4/9/24

X : Conti. RV with PDF f_X

$Y = u(X)$

Q: Dist of Y ?

$F_Y(y) = \mathbb{P}(Y \le y) = \mathbb{P}(u(X) \le y)$

Suppose u is strictly increasing

$u^{-1}(u(x)) = x$

$F_Y(y) = \mathbb{P}(X \le u^{-1}(y)) = F_X(u^{-1}(y))$ $\downarrow$ take derivative

$f_Y(y) = f_X(u^{-1}(y)) \cdot (u^{-1}(y))'$

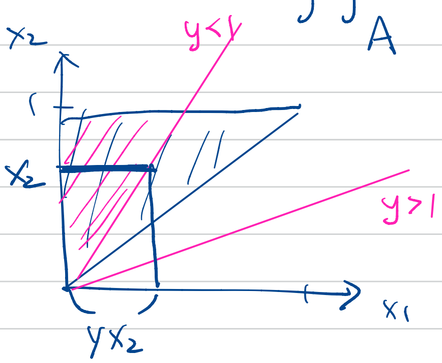
Example X_1, X_2 with joint PDF

$f(x_1, x_2) = 2$ for $0 < x_1 < x_2 < 1$

$Y = \dfrac{X_1}{X_2}$ Dist. of Y ?

$F_Y(y) = \mathbb{P}\left(\dfrac{X_1}{X_2} \le y\right) = \mathbb{P}(X_1 \le y X_2)$

$= \displaystyle\iint_A f(x_1, x_2)\, dx_1\, dx_2$



$x_1 = y \cdot x_2$

$x_2 = \left(\dfrac{1}{y}\right) \cdot x_1$ ↙ slope

If $y > 1$, $F_Y(y) = 1$

If $0 < y < 1$, $F_Y(y) = \displaystyle\int_0^1 \int_0^{y \cdot x_2} 2\, dx_1\, dx_2 = \int_0^1 2y \cdot x_2\, dx_2 = y$

$f_Y(y) = 1$ for $0 < y < 1$.

Q: X_1, X_2 as above

$Y_1 = \frac{X_1}{X_2}$, $Y_2 = X_2$

What is joint PDF of Y_1, Y_2

Transformations of Two Random Variables

If X_1 and X_2 are two continuous-type random variables with joint PDF $f(x_1, x_2)$.

Let $Y_1 = u_1(X_1, X_2)$, $Y_2 = u_2(X_1, X_2)$.

If $X_1 = v_1(Y_1, Y_2)$, $X_2 = v_2(Y_1, Y_2)$, then the joint PDF of Y_1 and Y_2 is

$$f_{Y_1,Y_2} = |J| f_{X_1,X_2}(v_1(y_1, y_2), v_2(y_1, y_2))$$

where J is the Jacobian given by

$$J := \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix}. = \begin{vmatrix} \frac{\partial v_1}{\partial y_1} & \frac{\partial v_1}{\partial y_2} \\ \frac{\partial v_2}{\partial y_1} & \frac{\partial v_2}{\partial y_2} \end{vmatrix}$$

X_1, X_2 with $f_{X_1,X_2}(x_1, x_2) = 2$ for $0 < x_1 < x_2 < 1$

$Y_1 = \frac{X_1}{X_2} = u_1(X_1, X_2)$ $\quad Y_2 = X_2 = u_2(X_1, X_2)$

$X_1 = Y_1 \cdot Y_2 = v_1(Y_1, Y_2)$ $\quad X_2 = Y_2 = v_2(Y_1, Y_2)$

$f_{Y_1,Y_2}(y_1, y_2) = f_{X_1,X_2}(y_1 y_2, y_2) \cdot |J|$

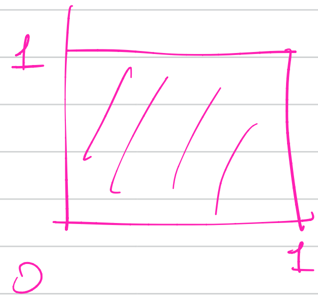
$|J| = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} \underset{\uparrow}{=} | y_2 \cdot 1 - y_1 \cdot 0 | = |y_2|$

Determinant
in Absolute Value

$$f_{Y_1,Y_2}(y_1,y_2) = 2y_2 \qquad 0 < \overset{y_1 y_2}{\overset{\|}{x_1}} < \overset{y_2}{\overset{\|}{x_2}} < 1$$

for $\underline{0 < y_1 y_2 < y_2 < 1}$

$\Rightarrow \begin{cases} 0 < y_1 < 1 \\ 0 < y_2 < 1 \end{cases}$



$$\Rightarrow f_{Y_1,Y_2}(y_1,y_2) = \underset{f_{Y_1}(y_1)}{\underset{\|}{(1)}} \cdot \underset{f_{Y_2}(y_2)}{\underset{\|}{(2y_2)}}$$

Y_1, Y_2 indep.

Transformations of Two Random Variables

Example

Let X_1 and X_2 have the joint PDF

$$f(x_1, x_2) = 2, \qquad 0 < x_1 < x_2 < 1.$$

Find the joint PDF of $Y_1 = \dfrac{X_1}{X_2}$ and $Y_2 = X_2$.

Exercise

Let X_1 and X_2 be independent random variables, each with PDF

$$f(x) = e^{-x}, \quad 0 < x < \infty.$$

Find the joint pdf of $Y_1 = X_1 - X_2$ and $Y_2 = X_1 + X_2$.

Section 3.
Several Independent Random Variables

Independent random variables

Recall that X_1 and X_2 are independent if

$$\mathbb{P}(X_1 \in A, X_2 \in B) = \mathbb{P}(X_1 \in A)\mathbb{P}(X_2 \in B)$$

for all A, B.

In particular, if X_1 and X_2 have PDFs, then $f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$.

A, B, C mutually indep.

$\mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$

$\mathbb{P}(B \cap C) = \mathbb{P}(B) \mathbb{P}(C)$

$\mathbb{P}(C \cap A) = \mathbb{P}(C) \mathbb{P}(A)$

$\mathbb{P}(A \cap B \cap C) = \mathbb{P}(A) \mathbb{P}(B) \mathbb{P}(C)$

Independent random variables

Definition

In general, we say $X_1, X_2, \cdots, X_n$ are independent if $\{X_1 \in A_1\}, \{X_2 \in A_2\}, \cdots, \{X_n \in A_n\}$ are mutually independent, for any choice of $A_1, A_2, \cdots, A_n$.

In particular, if $X_1, X_2, \cdots, X_n$ has PDFs, then the joint PDF is the product.

If $X_1, X_2, \cdots, X_n$ are independent and have the same distribution,

we say they are i.i.d. (independent and identically distributed) or a random sample of size n from that common distribution.

Independent random variables

Note : $X \sim Exp(\lambda)$, $\mathbb{P}(X > t) = e^{-\lambda t}$

$X_1, X_2, X_3 \sim Exp(1)$ i.i.d.

Example

Let X_1, X_2, X_3 be a random sample from a distribution with PDF

$$f(x) = e^{-x}, \quad 0 < x < \infty.$$

Find $\mathbb{P}(0 < X_1 < 1, 2 < X_2 < 4, 3 < X_3 < 7)$.

$\xrightarrow{\text{indep.}}$

$= \mathbb{P}(0 < X_1 < 1)\,\mathbb{P}(2 < X_2 < 4)\,\mathbb{P}(3 < X_3 < 7)$

$= (\mathbb{P}(X_1 > 0) - \mathbb{P}(X_1 > 1))(\mathbb{P}(X_2 > 2) - \mathbb{P}(X_2 > 4))(\mathbb{P}(X_3 > 3) - \mathbb{P}(X_3 > 7))$

$= (1 - e^{-1})(e^{-2} - e^{-4})(e^{-3} - e^{-7})$

Expectation and Variance

Theorem

Let $X_1, X_2, \cdots, X_n$ be a sequence of random variables. Then,

$$\mathbb{E}[X_1 + X_2 + \cdots + X_n] = \mathbb{E}[X_1] + \mathbb{E}[X_2] + \cdots + \mathbb{E}[X_n].$$

If they are independent, then

$$\mathbb{E}[X_1 X_2 \cdots X_n] = \mathbb{E}[X_1]\mathbb{E}[X_2] \cdots \mathbb{E}[X_n]$$

and

$$\mathrm{Var}[X_1 + X_2 + \cdots + X_n] = \mathrm{Var}[X_1] + \mathrm{Var}[X_2] + \cdots + \mathrm{Var}[X_n].$$

Note $\mathrm{Var}(X_1 + X_2) = \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + 2\,\mathrm{Cov}(X_1, X_2)$

$$\left(\mathrm{Var}(X_1 + X_2 + \cdots + X_n) = \mathrm{Var}(X_1) + \cdots + \mathrm{Var}(X_n) + 2 \sum_{1 \le i < j \le n} \mathrm{Cov}(X_i, X_j) \right)$$

Exercise

$$\mathbb{P}(X_1 > 4)^3 = (1 - \mathbb{P}(X_1 \leq 4))^3 = (1 - F(4))^3$$

$$\mathbb{P}(X_1 > k) = (1-p)^k \quad \Big| \quad \mathbb{P}(\underset{\underset{\text{Exp}(\lambda)}{\wr}}{X} > t) = (e^{-\lambda})^t$$

Let X_1, X_2, X_3 be i.i.d. Geometric with $p = \frac{3}{4}$.

Let Y be the minimum of X_1, X_2, X_3.

Find $\mathbb{P}(Y > 4)$.

$$= \mathbb{P}(X_1 > 4, X_2 > 4, X_3 > 4)$$

indep

$$= \mathbb{P}(X_1 > 4)\, \mathbb{P}(X_2 > 4)\, \mathbb{P}(X_3 > 4)$$

the same dist

$$= \mathbb{P}(X_1 > 4)^3$$

$$= \left(\left(\frac{1}{4}\right)^4 \right)^3 = \left(\frac{1}{4}\right)^{12}.$$

Section 4.
The Moment-Generating Function Technique

Recall $M_X(t) = \mathbb{E}[e^{tX}]$

Fact $M_X(t) = M_Y(t)$ around $t=0$

$\Rightarrow$ X, Y have the dist.

$M_X(t) \approx M_Y(t)$

$\Rightarrow$ X, Y "almost same"

The Moment-Generating Function

Theorem

If $X_1, X_2, \cdots, X_n$ are independent and have the MGFs $M_{X_i}(t)$, then the MGF of $Y = a_1 X_1 + \cdots a_n X_n$ is $M_Y(t) = M_{X_1}(a_1 t) \cdots M_{X_n}(a_n t)$.

Theorem

If $X_1, X_2, \cdots, X_n$ are i.i.d., then the MGF of $Y = X_1 + \cdots + X_n$ is $M_Y(t) = M_X(t)^n$. If $\overline{X} = \frac{X_1 + \cdots + X_n}{n}$, then the MGF is $M_{\overline{X}}(t) = M_X(\frac{t}{n})^n$.

$X_1, X_2, \cdots, X_n$ indep.

$$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$$

$$M_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t(a_1 X_1 + \cdots + a_n X_n)}]$$

$$= \mathbb{E}[e^{a_1 t X_1} \cdot e^{a_2 t X_2} \cdots e^{a_n t X_n}]$$

$$= \mathbb{E}[e^{a_1 t X_1}] \cdots \mathbb{E}[e^{a_n t X_n}] \quad \text{indep.}$$

$$= M_{X_1}(a_1 t) \cdots M_{X_n}(a_n t)$$

The Moment-Generating Function

Example

Let $X_1, X_2, \cdots, X_n$ be i.i.d. Bernoulli with p.

Let $Y = X_1 + \cdots + X_n$. Indep sum

Find the MGF of Y.

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t) \cdots M_{X_n}(t) \quad \text{(indep.)}$$

$$= \left(M_{X_1}(t)\right)^n \quad \text{(same dist.)}$$

$$M_{X_1}(t) = \mathbb{E}[e^{tX_1}] = e^{t \cdot 0} \cdot (1-p) + e^{t \cdot 1} \cdot p = 1 - p + pe^t$$

$$M_Y(t) = (1 - p + pe^t)^n$$

$$Y \sim Bin(n, p)$$

The Moment-Generating Function

Example

Let $X_1, X_2, \cdots, X_n$ be i.i.d. exponential with θ.

Let $Y = X_1 + \cdots + X_n$.

Find the MGF of Y.

Exercise

Let X_1, X_2, X_3 be independent Poisson with means $\underset{\lambda_1}{2}, \underset{\lambda_2}{1}, \underset{\lambda_3}{4}$.

Find the MGF of $Y = X_1 + X_2 + X_3$.

$$X_i \sim \text{Pois}(\lambda_i)$$

$$M_{X_i}(t) = e^{\lambda_i(e^t-1)}$$

$$M_Y(t) = M_{X_1}(t) \cdot M_{X_2}(t)\, M_{X_3}(t)$$

$$= e^{2(e^t-1)}\, e^{1\cdot(e^t-1)}\, e^{4(e^t-1)}$$

$$= e^{7(e^t-1)}$$

$$Y \sim \text{Pois}(7)$$

4/11/24

$X_1, X_2, \cdots, X_n$: i.i.d. (indep. & identically dist.)

$Y = a_1 X_1 + a_2 X_2 + \cdots + a_n X_n$

$M_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t\cdot(a_1X_1 + a_2X_2 + \cdots + a_nX_n)}]$

$= \mathbb{E}[e^{a_1 t X_1} e^{a_2 t X_2} \cdots e^{a_n t X_n}]$

$= \mathbb{E}[e^{a_1 t X_1}] \cdot \mathbb{E}[e^{a_2 t X_2}] \cdots \mathbb{E}[e^{a_n t X_n}]$ indep.

$= M_{X_1}(a_1 t)\, M_{X_2}(a_2 t) \cdots M_{X_n}(a_n t)$

Fact

(i) If $M_X(t) = M_Y(t) \Rightarrow X, Y$ have the same dist.

(ii) $M_X(t) \approx M_Y(t) \Rightarrow X, Y$ have the almost same dist.

Suppose $X_1, \cdots, X_n$ i.i.d. $\mathbb{E}[X_1] = 0 = \cdots = \mathbb{E}[X_n]$

$Y = a(X_1 + \cdots + X_n) \Rightarrow M_Y(t) = (M_X(at))^n$

$M_X(0) = 1, \quad M'_X(0) = \mathbb{E}[X] = 0, \quad M''_X(0) = \mathbb{E}[X^2] = Var(X) = \sigma^2$

$M_X(t) \approx M_X(0) + M'_X(0)\cdot\frac{t}{1!} + M''_X(0)\cdot\frac{t^2}{2!}$

when t is small, Taylor Expansion

$M_X(at) \approx 1 + \frac{\sigma^2 t^2}{2}\cdot a^2$

$M_Y(t) = (M_X(at))^n \approx \left(1 + \frac{\sigma^2 t^2}{2}\cdot a^2\right)^n = \left(1 + \left(\frac{\sigma^2 t^2}{2}\right)\cdot\frac{1}{n}\right)^n$

$a = \frac{1}{\sqrt{n}}$ for large n. $\to e^{\frac{\sigma^2 t^2}{2}}$ as $n \to \infty$

$\left(\lim_{n\to\infty}\left(1 + \frac{1}{n}\right)^n = e\right)$

If $Y = \frac{1}{\sqrt{n}}(X_1 + \cdots + X_n)$, then $M_Y(t) \longrightarrow e^{\frac{\sigma^2 t^2}{2}}$

Y converges to $N(0, \sigma^2)$ Central Limit Theorem

Section 6.
The Central Limit Theorem

The Central Limit Theorem

Let $X_1, X_2, \cdots, X_n$ be i.i.d. with common distribution X.

Let $\mathbb{E}[X] = \mu$ and $\mathrm{Var}(X) = \sigma^2$.

Let $\overline{X} = \frac{X_1 + \cdots + X_n}{n}$, then (Sample mean)

$\mathbb{E}[\overline{X}] = \mu$

$\mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}$

Let $W = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$, then

$$W = \frac{\overline{X} - \underbrace{\mathbb{E}[\overline{X}]}_{=\mu}}{\sqrt{\mathrm{Var}(\overline{X})}}$$

$\mathbb{E}[W] = 0$

$\mathrm{Var}(W) = 1$

$$\mathbb{E}[\overline{X}] = \mathbb{E}\left[\frac{1}{n}\cdot(X_1 + \cdots + X_n)\right] = \frac{1}{n}\mathbb{E}[X_1 + \cdots + X_n]$$

$$= \frac{1}{n}\Big(\underbrace{\mathbb{E}[X_1]}_{=\mu} + \cdots + \underbrace{\mathbb{E}[X_n]}_{=\mu}\Big) = \frac{1}{n}\cdot n\cdot\mu = \mu.$$

$$\mathrm{Var}(\overline{X}) = \mathrm{Var}\left(\frac{1}{n}(X_1 + \cdots + X_n)\right) = \frac{1}{n^2}\mathrm{Var}(X_1 + \cdots + X_n)$$

$$\overset{?}{=} \frac{1}{n^2}\Big(\underbrace{\mathrm{Var}(X_1)}_{=\sigma^2} + \cdots + \underbrace{\mathrm{Var}(X_n)}_{=\sigma^2}\Big) = \frac{1}{n^2}\cdot n\cdot\sigma^2 = \frac{\sigma^2}{n}$$

$X_1, \ldots, X_n$ i.i.d.

$\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$

The Central Limit Theorem

Theorem

If μ and σ^2 are finite, then the distribution of $W = \dfrac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ converges to that of the standard normal distribution $N(0,1)$ as $n \to \infty$.

The convergence is in the following sense: If n is large, for the standard normal Z,

$$\mathbb{P}(W \leq x) \approx \mathbb{P}(Z \leq x) =: \Phi(x) = \int_{\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{|y|^2}{2}}\, dy.$$

CDF of W $\rightarrow$ CDF of Standard Normal

Convergence in distribution.

The Central Limit Theorem

$X_1, X_2, \cdots, X_{25}$ i.i.d.

$\overline{X} = \frac{1}{25}(X_1 + X_2 + \cdots + X_n)$

Example

Let $\overline{X}$ be the mean of a random sample of $n = 25$ currents (in milliamperes) in a strip of wire in which each measurement has a mean of 15 and a variance of 4.

Find the approximate probability $\mathbb{P}(14.4 < \overline{X} < 15.6)$.

$$\mathbb{E}[X_1] = \cdots = \mathbb{E}[X_n] = 15 = \mu$$

$$Var(X_1) = \cdots = Var(X_n) = 4 = \sigma^2$$

$$n = 25$$

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - 15}{0.4}$$

Converges to $N(0,1)$ in dist

$(W \Rightarrow N(0,1))$

$$\mathbb{P}(14.4 < \overline{X} < 15.6)$$

$$= \mathbb{P}\left(\frac{14.4-\mu}{\sigma/\sqrt{n}} < \frac{\overline{X}-\mu}{\sigma/\sqrt{n}} < \frac{15.6-\mu}{\sigma/\sqrt{n}}\right) = \mathbb{P}\left(\frac{14.4-15}{0.4} < W < \frac{15.6-15}{0.4}\right)$$

$$= \mathbb{P}(-1.5 < W < 1.5) = \mathbb{P}(W < 1.5) - \mathbb{P}(W < -1.5)$$

$$\approx \mathbb{P}(Z < 1.5) - \mathbb{P}(Z < -1.5) = \Phi(1.5) - (1 - \Phi(1.5))$$

(By CLT)

$$= 2\Phi(1.5) - 1$$

The Central Limit Theorem

Example

Let $\overline{X}$ denote the mean of a random sample of size 25 from the distribution whose PDF is $f(x) = \frac{x^3}{4}$, $0 < x < 2$.

Find the approximate probability $\mathbb{P}(1.5 \leq \overline{X} \leq 1.65)$.

Exercise

Let X equal the maximal oxygen intake of a human on a treadmill, where the measurements are in milliliters of oxygen per minute per kilogram of weight.

Assume that, for a particular population, the mean of X is $\mu = 54.030$ and the standard deviation is $\boxed{\sigma = 5.8.}$

Let $\overline{X}$ be the sample mean of a random sample of size $n = 47$.

Find $\mathbb{P}(52.761 \leq \overline{X} \leq 54.453)$, approximately.

$$\overline{X} = \frac{1}{47}(X_1 + X_2 + \cdots + X_n)$$

$$\mu = 54.030$$

$$\sigma = 5.8$$

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

$$P(52.761 \leq \overline{X} \leq 54.453)$$

$$= P\left(\underbrace{\frac{52.761 - 54.030}{\frac{5.8}{\sqrt{47}}}}_{= -1.5} \leq W \leq \underbrace{\frac{54.453 - 54.030}{\frac{5.8}{\sqrt{47}}}}_{0.5}\right)$$

$$\approx P(-1.5 \leq Z \leq 0.5) = \Phi(0.5) - \Phi(-1.5)$$

By CLT

$$= \Phi(0.5) - (1 - \Phi(1.5))$$

$$= \Phi(0.5) + \Phi(1.5) - 1$$

Central Limit Theorem

$X_1, X_2, \cdots, X_n$ i.i.d. (indep. & the same dist.)

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

$\overline{X} = \frac{S_n}{n}$

Standard Normal

$$Y := \frac{\overline{X} - \mathbb{E}[\overline{X}]}{\sqrt{Var(\overline{X})}} = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} \Rightarrow N(0,1) \quad \text{as} \quad n \to \infty$$

Convergence in distribution

$$\mathbb{P}(Y \leq x) \longrightarrow P(Z \leq x) = \Phi(x)$$

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{Var(S_n)}} = \frac{S_n - n\cdot\mu}{\sqrt{n\cdot\sigma^2}} = \frac{\frac{S_n}{n} - \mu}{\sqrt{\sigma^2/n}} = Y \Rightarrow N(0,1)$$

Normal Approximation to Binomial

$Y \sim Bin(n, p)$

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

$X_1, \cdots, X_n$: i.i.d. $Ber(p)$

$$\frac{Y - \mathbb{E}[Y]}{\sqrt{Var(Y)}} = \frac{Y - np}{\sqrt{n\cdot p\cdot(1-p)}} \Rightarrow N(0,1) \quad \text{as} \quad n \to \infty$$

Poisson Approximation

$Bin(n, p) \approx Pois(\lambda)$ when p is small, n large

$np = \lambda$

# of Heads

$X \sim Bin(100, \frac{99}{100})$

# of Tails

$Y \sim Bin(100, \frac{1}{100})$

$X + Y = 100$

$$\mathbb{P}(\underline{X} \geq 50) = \mathbb{P}(100 - Y \geq 50)$$

$$= \mathbb{P}(\underline{Y \leq 50}) \approx \mathbb{P}(\cdots)$$

Section 7.
Approximations for Discrete Distributions

Normal approximation to Binomial Distribution

Theorem

Let X be a binomial random variable with parameter n and p. If n is large enough (usually, $np \geq 5$ and $n(1-p) \geq 5$), then X is approximately a normal distribution with mean np and variance $np(1-p)$.

$$\frac{X - np}{\sqrt{np(1-p)}} = \frac{X - \mathbb{E}[X]}{\sqrt{Var(X)}} \Rightarrow N(0,1)$$

Normal approximation to Binomial Distribution

Example

Let Y be Bin$(25, \frac{1}{2})$. Find the approximate probability $\mathbb{P}(12 \leq Y < 15)$ using the central limit theorem.

By CLT, $\dfrac{Y - E[Y]}{\sqrt{Var(Y)}} = \dfrac{Y - 12.5}{\sqrt{25/4}} \Rightarrow N(0,1)$

$$\mathbb{P}(12 \leq Y < 15) = \mathbb{P}\left(\frac{12-12.5}{2.5} \leq \frac{Y-12.5}{2.5} < \frac{15-12.5}{2.5}\right)$$

$$\approx \mathbb{P}(-0.2 \leq Z < 1)$$

$$= \Phi(1) - \Phi(-0.2) = \Phi(1) + \Phi(0.2) - 1.$$

$$\mathbb{P}(Y = 12) = \mathbb{P}(11.5 < Y < 12.5) \approx \mathbb{P}\left(\frac{11.5-12.5}{2.5} < Z < \frac{12.5-12.5}{2.5}\right)$$

half unit correction

mid-point correction

$$\mathbb{P}(\underline{12 \leq Y < 15}) = \mathbb{P}(Y = 12, 13, 14) = \mathbb{P}(Y=12) + \mathbb{P}(Y=13) + \mathbb{P}(Y=14)$$

$$= \mathbb{P}(\underline{11.5 \leq Y \leq 14.5}) \approx \mathbb{P}\left(\frac{11.5-12.5}{2.5} \leq Z \leq \frac{14.5-12.5}{2.5}\right)$$

$X \sim \text{Pois}(\underset{\overset{\shortparallel}{\lambda}}{n})$, n large

$X = X_1 + X_2 + \cdots + X_n$ $\quad X_1, \ldots, X_n$ i.i.d. Pois(1)

$$\frac{X - E(X)}{\sqrt{Var(X)}} = \frac{X - n}{\sqrt{n}} = \frac{X - \lambda}{\sqrt{\lambda}} \Rightarrow N(0, 1)$$

Normal approximation to Poisson Distribution

Theorem

Let X be a Poisson random variable with parameter λ. Then,

$$W := \frac{X - \lambda}{\sqrt{\lambda}}$$

converges to $N(0, 1)$ in distribution as $\lambda \to \infty$.

Normal approximation to Poisson Distribution

Example

Let $X_1, X_2, \cdots, X_{30}$ be a random sample of size 30 from a Poison distribution with a mean of $\frac{2}{3}$. Approximate the probability

$\lambda =$

$$\mathbb{P}\left(21 \leq \sum_{i=1}^{30} X_i \leq 27\right).$$

$$Y = \sum_{i=1}^{30} X_i \sim \mathrm{Pois}\left(\frac{2}{3} \cdot 30\right) = \mathrm{Pois}(20)$$

$$\frac{Y - 20}{\sqrt{20}} \approx N(0,1)$$

$$\mathbb{P}(21 \leq Y \leq 27) \approx \mathbb{P}\left(\frac{21-20}{\sqrt{20}} \leq Z \leq \frac{27-20}{\sqrt{20}}\right)$$

without half unit correction

$$\approx \mathbb{P}\left(\frac{20.5-20}{\sqrt{20}} \leq Z \leq \frac{27.5-20}{\sqrt{20}}\right)$$ with half unit correction.

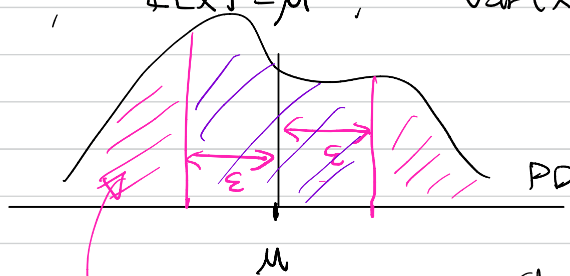
$$\mathbb{P}(21 \leq Y \leq 27) = \mathbb{P}(Y=21) + \mathbb{P}(Y=22) + \cdots + \mathbb{P}(Y=27)$$

Section 8.
Chebyshev's Inequality and Convergence in Probability

$\Rightarrow$ Law of Large Numbers

Recall

X , $\mathbb{E}[X] = \mu$, $Var(X) = \sigma^2$



Chebyshev's inequality

$$P(|X-\mu| \geq \varepsilon) \leq \frac{\mathbb{E}[|X-\mu|^2]}{\varepsilon^2} = \frac{\sigma^2}{\varepsilon^2}$$

$$P(|X-\mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

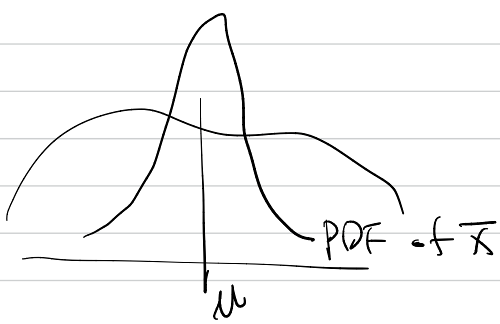
$X_1, X_2, \cdots, X_n$: i.i.d. $\quad \mathbb{E}[X_1] = \mu$

$Var(X_1) = \sigma^2$

$\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n)$ $\quad \mathbb{E}[\bar{X}] = \mu$

$Var(\bar{X}) = \frac{\sigma^2}{n}$

for large n



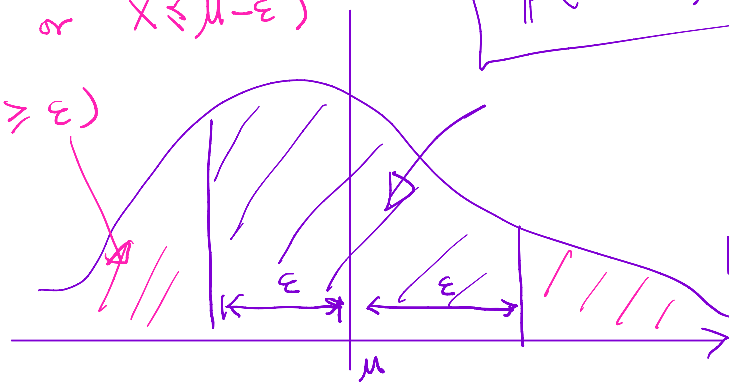
$$P\left(|\bar{X}-\mu| > \frac{1}{100}\right) \leq \frac{\mathbb{E}[|\bar{X}-\mu|^2]}{\left(\frac{1}{100}\right)^2} = \frac{\sigma^2}{n} \cdot (100)^2$$

Chebyshev

$$\boxed{\mathbb{P}(|X-\mu| < \varepsilon) \geq 1 - \frac{\sigma^2}{\varepsilon^2}}$$

μ σ

$$\mathbb{P}(X \geq \mu+\varepsilon \text{ or } X \leq \mu-\varepsilon)$$
$$= \mathbb{P}(|X-\mu| \geq \varepsilon)$$
$$\leq \frac{\sigma^2}{\varepsilon^2}$$



Chebyshev's Inequality

Theorem

If the random variable X has a mean μ and variance σ^2, then for every $k \geq 1$,

$$\mathbb{P}(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

In particular $\varepsilon = k\sigma$, then

$$\mathbb{P}(|X-\mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

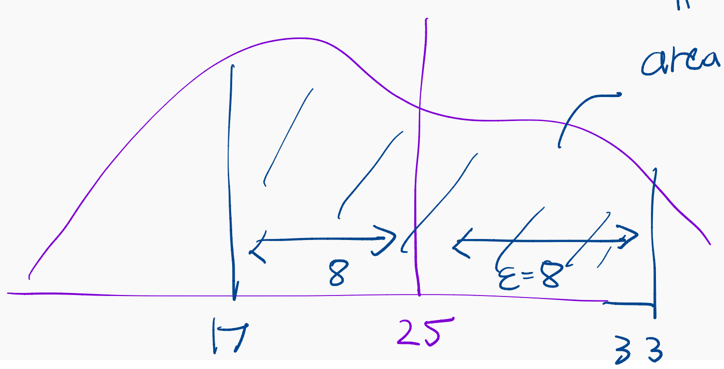
Chebyshev's Inequality

Example

Suppose X has a mean of 25 and a variance of 16.

Find the lower bound of $\mathbb{P}(17 < X < 33)$.

$$\mathbb{P}(17 < X < 33) = \mathbb{P}(|X - 25| < 8) = \text{area} \geq 1 - \frac{\sigma^2}{\varepsilon^2} = 1 - \frac{16}{8^2} = \frac{3}{4}.$$



The Law of Large Numbers

Definition

We say a sequence of random variables X_n converges to a random variable X in probability if for every $\varepsilon > 0$,

$$\lim_{n\to\infty} \mathbb{P}(|X_n - X| > \varepsilon) = 0.$$

The Law of Large Numbers

Law of Large Numbers

Theorem

Let $X_1, X_2, \cdots, X_n$ be i.i.d. with common distribution X.

Let $\mathbb{E}[X] = \mu$ and $\text{Var}(X) = \sigma^2$.

Then, $\overline{X}$ converges to μ in probability.

For any $\varepsilon > 0$

$$\mathbb{P}(|\overline{X} - \mu| \geq \varepsilon) \longrightarrow 0 \quad \text{as} \quad n \to \infty$$

Exercise

If X is a random variable with mean 33 and variance 16, use Chebyshev's inequality to find

1. A lower bound for $\mathbb{P}(23 < X < 43)$.
2. An upper bound for $\mathbb{P}(|X - 33| \geq 14)$.