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.

Examples : $Y = X^2$

$Y = e^X$

$\vdots$

"Find Dist. of a New RV $\Rightarrow$ Look at $\boxed{\text{CDF}}$"

Functions of One Random Variable

Example

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

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

$X = -2, -1, 0, 1, \cdots, 5$ equally likely.

$f_X(k) = \text{constant} = \frac{1}{8}$ for $k = \underbrace{-2, -1, \cdots, 5}_{8}$

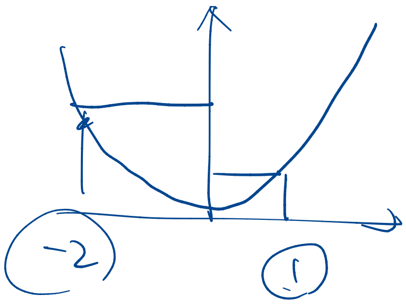
$Y = X^2$

$= (-2)^2, (-1)^2, 0^2, 1^2, \cdots, 5^2$

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

Discrete RV → a New from given one

Think PMF Directly.



$-2 \leq 1$

$\cdot \; P(X \leq Y) \overset{?}{=} P(X^2 \leq Y^2)$

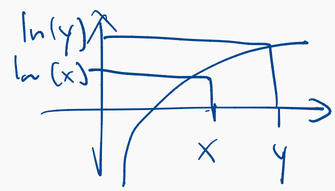
$u(X) \quad u(Y)$

$u(t) = t^2$

$\cdot \; P(0 \leq X \leq Y) = P(X^2 \leq Y^2)$

$\cdot \; P(X \leq Y < 0) = P(X^2 \geq Y^2)$

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 $Y = e^X$. $Y > 0$

$F_Y(y) = P(Y \leq y) = 0 \quad$ if $\; y \leq 0$

For $y > 0$,

$$F_Y(y) = P(Y \leq y) = P(e^X \leq y)$$

$$= P(\ln(e^X) \leq \ln y)$$

$$= P(X \leq \ln y) = F_X(\ln y)$$

because $\ln(t)$ is increasing

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

Chain Rule

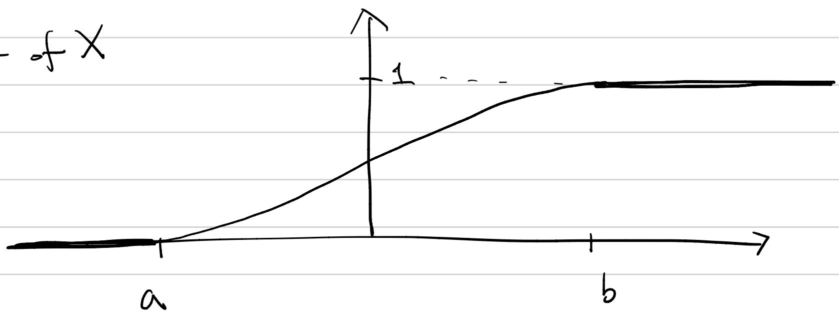
$$= \frac{1}{y} \frac{1}{\Gamma(\alpha)\theta^\alpha} (\ln y)^{\alpha-1} e^{-\frac{1}{\theta}(\ln y)}$$

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

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

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

CDF of X



$F_X = F$.

$\downarrow$

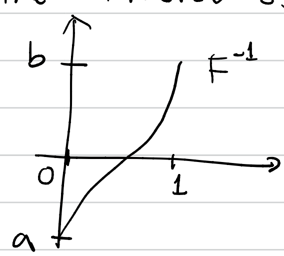
F^{-1}

$\uparrow$

The inverse of F on (a,b)

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

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



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

$Y = F^{-1}(U)$

Q: Dist. of Y. ?

$e^{x} \le y \rightarrow \ln(e^{x}) \le \ln y)$

$$F_Y(y) = P(Y \le y) = P(F^{-1}(U) \le y)$$

because F is increasing

$$= P(\underbrace{F(F^{-1}(U))}_{=U} \le F(y))$$

$$= P(U \le F(y))$$

$$= F_X(y) \Rightarrow X, Y \text{ have the same distribution.}$$

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, $X = F^{-1}(Y)$.

Application

WANT : Choose 1 sample from $\text{Exp}(2)$

Sample 1 number from $U(0,1)$, say x

$\underbrace{F^{-1}(x)}_{\sim \text{Exp}(2)}$, F is CDF of $\text{Exp}(2)$

$F(t) = 1 - e^{-2t} =$

$F^{-1}(t) = -\frac{1}{2}\ln(1-t)$

$\left(\begin{array}{l} t = 1 - e^{-2F^{-1}(t)} \\ e^{-2F^{-1}(t)} = 1 - t \end{array} \right)$

$f_X(x) = 3(1-x)^2 \rightarrow F_X(x) = 1 - (1-x)^3$

$F(X) = 1 - \underbrace{(1-X)^3}_{=Y} = 1-Y$

Change of Variables

$0 < (1-x)^3 < 1$

$0 < 1-x < 1$

$\Uparrow$

$0 < X < 1$

Example

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

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

For $0 < y < 1$

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

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

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

$f_Y(y) = 1$ for $y \in (0,1)$

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

$f_Y(y) = \frac{d}{dy}(1 - F_X(1-y^{\frac{1}{3}}))$

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

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

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

Exer: $X \sim \text{Exp}(1)$, Dist. of e^{-X}.

In general,

$\underline{\underline{X}}$

Define $\quad Y = u(X)$

Assume $\quad u$ is monotone $\left\langle \begin{array}{l} \text{increasing} \\ \text{decreasing} \end{array}\right.$

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

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

$$= \begin{cases} \mathbb{P}(X \leq u^{-1}(y)) = F_X(u^{-1}(y)) & \text{if } u \text{ increasing} \\ \mathbb{P}(X \geq u^{-1}(y)) \overset{c}{=} 1 - F_X(u^{-1}(y)) & \text{if } u \text{ decreasing} \end{cases}$$

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

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

Exercise

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

Find the PDF of $Y = X^2$. $= u(X)$

$u(t) = t^2$, $\underline{\underline{0<t<1}}$

monotone? Yes.

$u^{-1}(t) = \sqrt{t}$

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

$$= f_X(\sqrt{y}) \cdot \frac{1}{2\sqrt{y}}$$

$$= 4 \cdot y^{\frac{3}{2}} \cdot \frac{1}{2} \cdot y^{-\frac{1}{2}} = 2y.$$

Section 2.
Transformations of Two Random Variables

4/9/24

- X : a continuous RV with PDF $f_X(x)$

$Y = u(X)$ What is the PDF of Y ?

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

if u is strictly increasing, there exists the inverse of u.

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

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

↓ take derivative in y

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

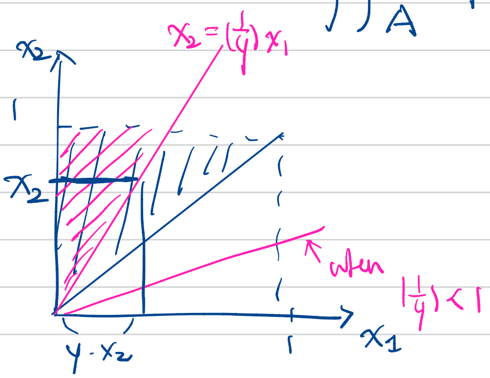
<u>Example</u> X_1, X_2 : continuous RV with joint PDF

$$f(x_1, x_2) = 2 \quad \text{for} \quad 0 < x_1 < x_2 < 1$$

Define $Y = \frac{X_1}{X_2}$. What is the PDF of Y?

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

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



$A = \{(x_1, x_2) : x_1 \le y \cdot x_2\}$

$x_2 \ge \left(\frac{1}{y}\right) \cdot x_1$ ← slope

If $y \ge 1$, the $F_Y(y) = 1$

For $0 < y < 1$, $F_Y(y) = \int_0^1 \int_0^{yx_2} 2\, dx_1\, dx_2 = \int_0^1 2 \cdot x_2 \cdot y\, dx_2$

$= y$

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

Q: X_1, X_2 $\Rightarrow$ $Y_1 = \frac{X_1}{X_2}$, $Y_2 = X_2$

the 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 $f(x_1, x_2)$.

$= x_1/x_2$ $\quad$ $= x_2$

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 $(v_1, v_2)'$ & matrix

determinant in absolute value

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

Example

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

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

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

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

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

$$= 2 \cdot y_2 \quad \text{for} \quad 0 < y_1 \cdot y_2 < y_2 < 1$$

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 $f_{X_1,X_2}(x_1, x_2) = f_{X_1}(x_1)f_{X_2}(x_2)$.

$\{X_1 \in A_1\}$, $\{X_2 \in A_2\}$, $\{X_3 \in A_3\}$ mutually indep.

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2) = \mathbb{P}(X_1 \in A_1) \cdot \mathbb{P}(X_2 \in A_2)$$

$$\mathbb{P}(X_2 \in A_2, X_3 \in A_3) = \mathbb{P}(X_2 \in A_2)\,\mathbb{P}(X_3 \in A_3)$$

$$\mathbb{P}(X_3 \in A_3, X_1 \in A_1) = \mathbb{P}(X_3 \in A_3)\,\mathbb{P}(X_1 \in A_1)$$

$$\mathbb{P}(X_1 \in A_1, X_2 \in A_2, X_3 \in A_3) = \mathbb{P}(X_1 \in A_1)\,\mathbb{P}(X_2 \in A_2)\,\mathbb{P}(X_3 \in A_3)$$

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, (PDFs are the same)

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

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

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

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

$\left(= \int_0^1 e^{-x}\,dx \cdot \int_2^4 e^{-x}\,dx \cdot \int_3^7 e^{-x}\,dx = \right)$

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

- $\mathbb{P}(0 < X < 1) = F(1) - F(0)$ $\quad F(x) = 1 - e^{-x}$
- $X \sim \text{Exp}(\lambda)$ $\quad \mathbb{P}(X > t) = e^{-\lambda t}$

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

Expectation and Variance

Theorem

works even for dependent cases

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

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

$$Var(X_1 + X_2) = Var(X_1) + Var(X_2) + 2\,Cov(X_1, X_2)$$

Exercise

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

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

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

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

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

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

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

Section 4. The Moment-Generating Function Technique

Example X_1, X_2, X_3 $Y = \frac{(X_1+X_2+X_3)}{3}$

$= \frac{1}{3} \cdot X_1 + \frac{1}{3} X_2 + \frac{1}{3} \cdot X_3$

$\Rightarrow M_Y(t) = M_{X_1}(\frac{1}{3}t) \, M_{X_2}(\frac{1}{3}t) \, M_{X_3}(\frac{1}{3}t)$

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 $Y = a_1X_1 + \cdots a_nX_n$ is $M_Y(t) = M_{X_1}(a_1t)\cdots M_{X_n}(a_nt)$.

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 $\boxed{M_{\overline{X}}(t) = M_X(\frac{t}{n})^n}$.

$3(X_1 + \cdots + X_n) \Rightarrow M_Y(t) = M_{X_1}(3t)^n$

The Moment-Generating Function

Example

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

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

Find the MGF of Y.

$X_1 \sim Ber$

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

$$= 1 - p + e^t \cdot p$$

$$M_Y(t) = (M_{X_1}(t))^n$$

$$= \boxed{(1 - p + e^t \cdot p)^n}$$

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

The Moment-Generating Function

Example

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

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

Find the MGF of Y.

$$M_{X_1}(t) = \frac{1}{(1-\theta t)}$$

$$M_Y(t) = \frac{1}{(1-\theta t)^n} = M_{X_1}(t)^n = \frac{1}{(1-\frac{t}{\lambda})^n}$$

$$Y \sim \text{Gamma}(n, \lambda)$$

Exercise

$$\lambda_1 \quad \lambda_2 \quad \lambda_3$$

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

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

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

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

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

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

$$Y \sim \text{Pois}(2+1+4)$$

4/11/24

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

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

$M_Y(t) = \mathbb{E}[e^{tY}] = \mathbb{E}[e^{t\cdot(X_1+X_2+\cdots+X_n)}]$

$= \mathbb{E}[e^{tX_1} \cdot e^{tX_2} \cdots e^{tX_n}]$

$\Downarrow$ $X_1, \cdots, X_n$ indep

$= \mathbb{E}[e^{tX_1}]\,\mathbb{E}[e^{tX_2}] \cdots \mathbb{E}[e^{tX_n}]$

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

$\Downarrow$ $X_1, \cdots, X_n$ have the same distribution

$= (M_{X_1}(t))^n$

MGF Technique

If $Y = a_1X_1 + a_2X_2 + \cdots + a_nX_n$

$M_Y(t) = M_{X_1}(a_1t)\, M_{X_2}(a_2t) \cdots M_{X_n}(a_nt)$

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

$a = a_1 = a_2 = \cdots = a_n$, $Y = aX_1 + aX_2 + \cdots + aX_n$

$M_Y(t) = (M_X(at))^n$

$M_X(t)$

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

$M_X(t) \approx M_X(0) + M_X'(0)\cdot t + M_X''(0)\cdot\frac{t^2}{2} = 1 + \frac{\sigma^2t^2}{2}$

(Taylor Expansion)

$M_Y(t) = (\underline{M_X(at)})^n \approx \left(1 + \frac{\sigma^2a^2t^2}{2}\right)^n \xrightarrow[n\to\infty]{}$

$a^2 = \frac{1}{n}$ $a = \frac{1}{\sqrt{n}}$

$Y = \frac{1}{\sqrt{n}}\cdot(X_1 + \cdots + X_n)$

$\left(1 + \frac{1}{n}\cdot\left(\frac{\sigma^2t^2}{2}\right)\right)^n \longrightarrow e^{\frac{\sigma^2t^2}{2}}$

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

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

$$Y = \frac{1}{\sqrt{n}}(X_1+\dots+X_n) \Rightarrow N(0,\sigma^2)$$

$$M_Y(t) \approx e^{\frac{\sigma^2 t^2}{2}} \quad \text{for large } n.$$

Central Limit Theorem.

Section 6.
The Central Limit Theorem

Note

$$Y = \frac{X - \mathbb{E}[X]}{\sqrt{Var(X)}}$$

$$\mathbb{E}[Y] = \mathbb{E}\left[\frac{X - \mathbb{E}[X]}{\sqrt{Var(X)}}\right] = \frac{1}{\sqrt{Var(X)}}\left(\underbrace{\mathbb{E}[X - \mathbb{E}[X]]}_{= \mathbb{E}[X] - \mathbb{E}[X] = 0}\right)$$

$$Var(Y) = Var\left(\frac{X - \mathbb{E}[X]}{\sqrt{Var(X)}}\right) = \frac{Var(X - \mathbb{E}[X])}{Var(X)} = 1.$$

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

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

$\overline{X}$ = the sample mean.

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

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

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

$$W = \frac{\overline{X} - \mathbb{E}[\overline{X}]}{\sqrt{Var(\overline{X})}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

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

$\text{Var}(W) = 1.$

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

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

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

$$= \frac{1}{n^2} Var(X_1 + \cdots + X_n) = \frac{1}{n^2}\cdot\left(Var(X_1) + Var(X_2) + \cdots + Var(X_n)\right)$$

$$= \frac{1}{n^2}\cdot n\cdot\sigma^2 = \frac{\sigma^2}{n}.$$

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

$\overline{X} = \frac{1}{n}(X_1 + \cdots + X_n)$, $W = \frac{\overline{X} - E[\overline{X}]}{\sqrt{Var(\overline{X})}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$

The Central Limit Theorem

Theorem

Convergence in dist.

If μ and σ^2 are finite, then the distribution of $W = \frac{\overline{X} - \mu}{\frac{\sigma}{\sqrt{n}}}$ 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

CDF of Normal

The Central Limit Theorem

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \Rightarrow N(0,1) \quad \text{meaning that}$$

$$\mathbb{P}(W \le t) \approx \Phi(t) \quad (\text{CDF of Standard Normal})$$

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

$X_1, X_2, \dots, X_{25}$ i.i.d. $\mathbb{E}[X_i] = 15 = \mu$, $Var(X_i) = 4 = \sigma^2$

$$\overline{X} = \frac{1}{n}(X_1 + \dots + X_n) = \frac{1}{25}(X_1 + \dots + X_{25})$$

$\sigma = 2$

$\sqrt{n} = \sqrt{25} = 5$

$\frac{\sigma}{\sqrt{n}} = \frac{2}{5} = 0.4$

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

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

$$= \mathbb{P}\left(-\frac{0.6}{0.4} < W < \frac{0.6}{0.4}\right) = \mathbb{P}(-1.5 < W < 1.5)$$

$$\underset{\text{By } \underline{\text{CLT}}}{\approx} \mathbb{P}(-1.5 < Z < 1.5) = \Phi(1.5) - \Phi(-1.5)$$

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

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

$$\mathrm{Var}(X) = \mathbb{E}[X^2] - (\mathbb{E}[X])^2 = \frac{8}{3} - \left(\frac{8}{5}\right)^2 = 8\left(\frac{1}{3} - \frac{8}{25}\right) = \frac{8}{75}$$

The Central Limit Theorem

$$\mathbb{E}[X] = \int x f(x)\,dx = \int_0^2 x \cdot \frac{x^3}{4}\,dx = \left[\frac{1}{4}\cdot\frac{1}{5}\cdot x^5\right]_0^2 = \frac{32}{20} = \frac{8}{5}$$

$$\mathbb{E}[X^2] = \int x^2 f(x)\,dx = \int_0^2 x^2 \cdot \frac{x^3}{4}\,dx = \left[\frac{1}{4}\cdot\frac{1}{6}\cdot x^6\right]_0^2 = \frac{64}{24} = \frac{8}{3}$$

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 $\mathbb{P}(1.5 \leq \overline{X} \leq 1.65)$.

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

$\mu = \frac{8}{5}$ $\quad \sigma^2 = \frac{8}{75}$

$n = 25$

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

$$\frac{\sigma}{\sqrt{n}} = \sqrt{\frac{\sigma^2}{n}} = \sqrt{\frac{8}{75}\cdot\frac{1}{25}} = \frac{\sqrt{8/3}}{25}$$

$$\mathbb{P}(1.5 \leq \overline{X} \leq 1.65)$$

$$= \mathbb{P}\left(\frac{1.5 - 8/3}{\frac{\sqrt{8/3}}{25}} \leq W \leq \frac{1.65 - 8/3}{\frac{\sqrt{8/3}}{25}}\right)$$

$$\approx \Phi(\quad) - \Phi(\quad)$$

By CLT.

for large $n = 25$

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 $\sigma = 5.8$.

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

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

Central Limit Theorem

$X_1, X_2, \cdots, X_n$: i.i.d. (Independent , same $\overset{X}{\overset{\shortparallel}{\text{dist.}}}$)

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

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

$\bar{X} = \frac{S_n}{n}$: sample mean ($\mathbb{E}[\bar{X}] = \mu$, $Var(\bar{X}) = \frac{\sigma^2}{n}$)

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

Convergence in distribution

(meaning that $\mathbb{P}(W \leq x) \to \mathbb{P}(Z \leq x)$ as $n \to \infty$)

$Z \sim N(0,1)$

$$\frac{S_n - \mathbb{E}[S_n]}{\sqrt{Var(S_n)}} = \frac{(X_1 + \cdots + X_n) - n \cdot \mu}{\sqrt{n \cdot \sigma^2}} = W \Rightarrow N(0,1).$$

① $Y \sim Bin(n,p)$ $\quad Y = X_1 + X_2 + \cdots + X_n$

$X_1, \cdots, X_n$: i.i.d. $Ber(\frac{1}{p})$

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

$$\shortparallel$$

$$\frac{Y - n \cdot p}{\sqrt{np(1-p)}}$$

n is large

Normal Approximation to Binomial.

(Poisson Approximation : $Bin(n,p) \approx Pois(\lambda)$)

if n is large p is small

$np \approx \lambda$

② $Y \sim Pois(\overset{\lambda}{\overset{\shortparallel}{n}})$ $\quad Y = X_1 + X_2 + \cdots + X_n$

$X_1, \cdots, X_n$: i.i.d $Pois(1)$

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

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 $np(1-p)$.

$$\frac{X - np}{\sqrt{np(1-p)}} \approx N(0, 1)$$

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.

$n = 25$, $p = \frac{1}{2}$ $\mathbb{E}[Y] = 25 \cdot \frac{1}{2} = \frac{25}{2}$, $Var(Y) = 25 \cdot \frac{1}{2} \cdot (1 - \frac{1}{2})$
$= \frac{25}{4}$

$$\frac{Y - \frac{25}{2}}{\sqrt{25/4}} \Rightarrow N(0,1)$$

$$\mathbb{P}(12 \leq Y < 15) = \mathbb{P}\left(\frac{12 - \frac{25}{2}}{\frac{5}{2}} \leq \frac{Y - \frac{25}{2}}{\frac{5}{2}} < \frac{15 - \frac{25}{2}}{\frac{5}{2}}\right)$$

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

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

half unit correction
mid point correction

$$= \mathbb{P}(-0.4 < Z < 0)$$

$$\mathbb{P}(12 \leq Y < 15) = \mathbb{P}(Y = 12, 13, 14) = \mathbb{P}(11.5 < Y < 14.5)$$
$$\approx \mathbb{P}\left(\frac{11.5 - 12.5}{2.5} < Z < \frac{14.5 - 12.5}{2.5}\right)$$

Normal approximation to Poisson Distribution

Theorem

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

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

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

Normal approximation to Poisson Distribution

Example

i.i.d. Pois $\left(\frac{2}{3}\right)$

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

λ

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

$$\| $$

$$Y \sim \text{Pois}(20)$$

$$\frac{Y - \mathbb{E}[Y]}{\sqrt{\text{Var}(Y)}} = \frac{Y - 20}{\sqrt{20}} \Rightarrow N(0,1)$$

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

without half unit correction.

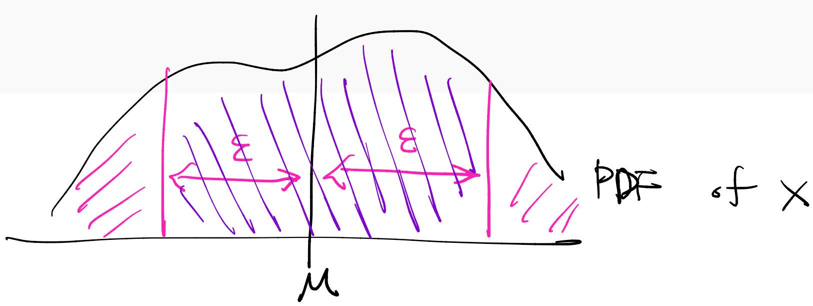
$$\mathbb{P}(21 \leq Y \leq 27) = \mathbb{P}(20.5 \leq Y \leq 27.5)$$

with half unit correction

$$\mathbb{P}(Y=21) + \mathbb{P}(Y=22) + \mathbb{P}(Y=23) + \cdots + \mathbb{P}(Y=27)$$

Section 8.
Chebyshev's Inequality and Convergence in Probability

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



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

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

Chebyshev's Inequality

Theorem

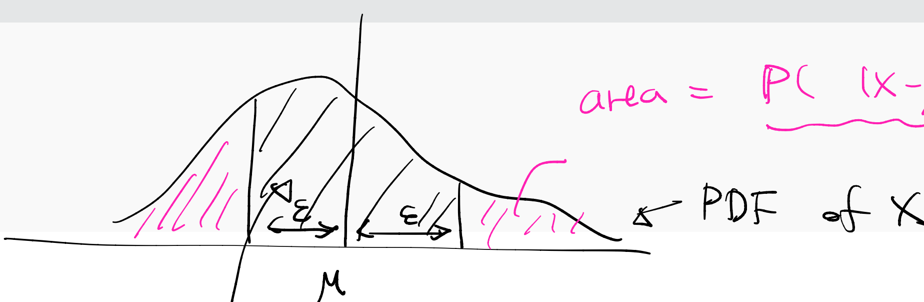
If the random variable X has a mean μ and variance σ^2, then for every $k \geq 1$,

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

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

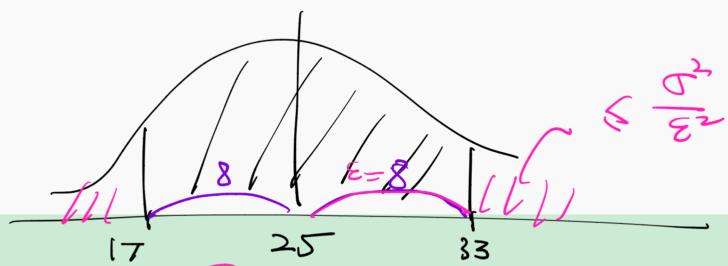
$$\mathbb{P}(|X - \mu| \geq k\sigma) \leq \frac{1}{k^2}.$$

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



area $= \mathbb{P}(\mu - \varepsilon < X < \mu + \varepsilon) = \mathbb{P}(\underbrace{|X - \mu| < \varepsilon})$

Chebyshev's Inequality



$\le \frac{\sigma^2}{\varepsilon^2}$

Example

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

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

$$= \mathbb{P}(|X-25| < 8) \geq 1 - \frac{\sigma^2}{\varepsilon^2}$$

$$= 1 - \frac{16}{8^2} = 1 - \frac{1}{4}$$

$$= \frac{3}{4}.$$

The Law of Large Numbers

(LLN)

Definition $X_1, X_2, X_3, \cdots,$

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

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

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

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

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

$\overline{X} \approx \mathbb{E}[X_1]$

Intuition.

$\mathbb{P}(|\overline{X} - \mu| > \frac{1}{100}) \leq \frac{\mathbb{E}[|\overline{X}-\mu|^2]}{(\frac{1}{100})^2} = \frac{Var(\overline{X})}{(\frac{1}{100})^2}$

empirical mean

Chebychev's Ineq.

$= \frac{\sigma^2}{n} \cdot (100)^2$

The 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 $\mathrm{Var}(X) = \sigma^2$.

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

relative frequency

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

Relative frequency $= \dfrac{X_1 + \cdots + X_n}{n} \xrightarrow{} \mathbb{E}[X_1] = p$ success prob.

LLN

Exercise

If X is a random variable with mean 3 and variance 16, use Chebyshev's inequality to find

1. A lower bound for $\mathbb{P}(23 < X < 43)$.
2. An upper bound for $\mathbb{P}(|X - 31| \geq 14)$.