Chapter 5. Distributions of Functions of Random Variables

Math 3215 Summer 2023

Georgia Institute of Technology

Section 1.
Functions of One Random Variable

Functions of One Random Variable

Let X be a random variable.

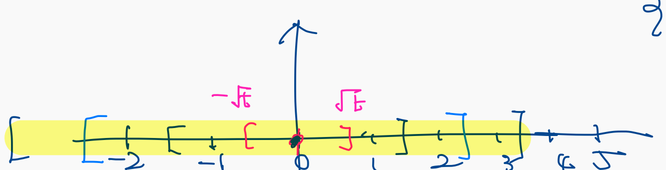
Define $Y = u(X)$ for some function u.

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

look at CDF

Functions of One Random Variable

$$\text{pmf of } X = f_X(x) = \begin{cases} \frac{1}{8}, & x = -2,-1,0,1,2,3,4,5 \quad \underline{\underline{8}} \\ 0, & \text{o.w.} \end{cases}$$



Example

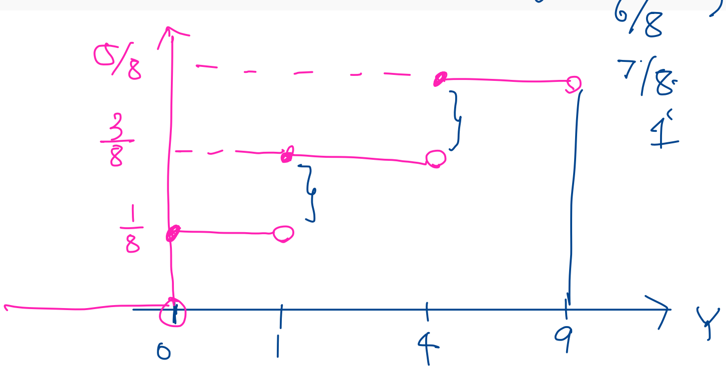
Let X have a discrete uniform distribution on the integers from $\underline{-2}$ to 5.

Find the distribution of $Y = X^2$. ≥ 0

$$P(\underline{\underline{Y}} \leq t) = P(\underline{X^2 \leq t}) = P(-\sqrt{t} \leq X \leq \sqrt{t})$$

$t \geq 0$

$$= \begin{cases} \frac{1}{8}, & 0 \leq t < 1 \\ \frac{3}{8}, & 1 \leq \sqrt{t} < 2 \rightarrow 1 \leq t < 4 \\ 5/8, & 2 \leq \sqrt{t} < 3 \rightarrow 4 \leq t < 9 \\ 6/8, & 3 \leq \sqrt{t} < 4 \rightarrow 9 \leq t < 16 \\ 7/8, & 4 \leq \sqrt{t} < 5 \rightarrow 16 \leq t < 25 \\ 1, & 5 \leq \sqrt{t} \longrightarrow t \geq 25 \end{cases}$$



$$P(\underline{\underline{Y}} = \underline{\underline{k}}) = P(X = \underline{\sqrt{k}} \text{ or } \underline{-\sqrt{k}})$$

$k = 0, 1, 4, 9, 16, 25$

$$f_Y(16) = f_Y(9) = f_Y(0) = \frac{1}{8}, \quad f_Y(1) = \frac{2}{8} = f_Y(4)$$

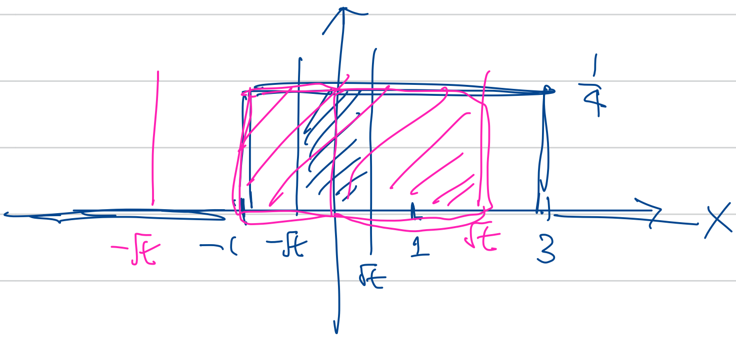
$$\| \quad f_Y(25)$$

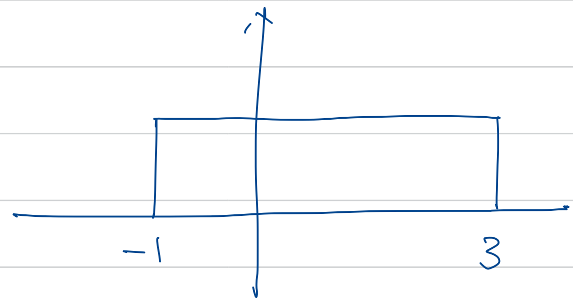
$X \sim \text{Unif}(-1,3)$ $\qquad Y = X^2 \geq 0$

$P(Y \leq t) = 0$ if $t < 0$

If $0 < t < 1$

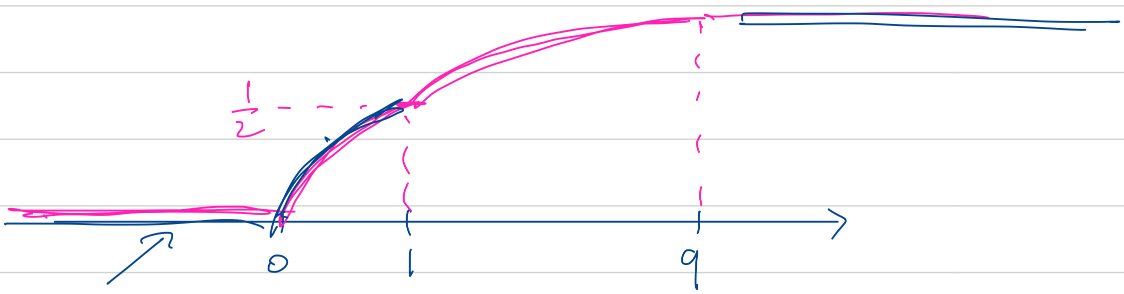
$P(Y \leq t) = P(-\sqrt{t} \leq X \leq \sqrt{t}) = 2 \cdot \sqrt{t} \cdot \frac{1}{4} = \frac{\sqrt{t}}{2}$





If $1 < \sqrt{t} < 3$, $P(Y \leq t) = \frac{1}{4} + \frac{1}{4}\sqrt{t}$

If $\sqrt{t} \geq 3$ $\quad P(Y \leq t) = 1$



$F_Y(t)$

$$f_Y(t) = \frac{d}{dt} F_Y(t) = \begin{cases} 0 & t \leq 0 \text{ or } t \geq 9 \\ \frac{1}{4\sqrt{t}} & 0 < t < 1 \\ \frac{1}{8\sqrt{t}} & 1 \leq t \leq 9 \end{cases}$$

CDF Technique

$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1} e^{-x} dx$, $\Gamma(n) = (n-1)!$

$\Gamma(\alpha) = (\alpha-1)\Gamma(\alpha-1)$

Example

Let X have a gamma distribution with pdf

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

$(\log = \ln = \log_e)$

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

$F_Y(t) = P(Y \leq t) = P(e^X \leq t) = \underbrace{P(X \leq \log t)} = F_X(\log t)$

$f_Y(t) = \frac{d}{dt} F_Y(t) = \frac{d}{dt}\left(F_X(\log t)\right)$

$\underset{\uparrow}{=} F_X'(\log t) \cdot (\log t)'$

Chain Rule

$= f_X(\log t) \cdot \frac{1}{t}$

$= \frac{1}{\Gamma(\alpha)\theta^\alpha} (\log t)^{\alpha-1} \cdot e^{-\frac{1}{\theta}\log t} \cdot \frac{1}{t}$.

$e^{-\frac{1}{\theta}\log t} = e^{\log\left(t^{-\frac{1}{\theta}}\right)} = t^{-\frac{1}{\theta}}$

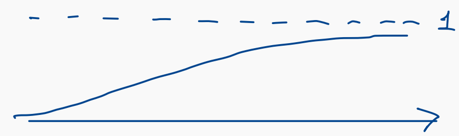
$= \frac{1}{\Gamma(\alpha)\theta^\alpha} \cdot (\log t)^{\alpha-1} \cdot t^{-\frac{1}{\theta}-1}$ //

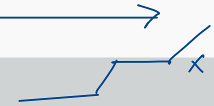
CDF Technique

CDF : ① non-decreasing

② $\lim_{x \to -\infty} F(x) = 0$

③ $\lim_{x \to \infty} F(x) = 1$





Theorem

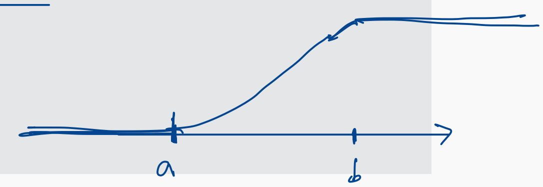
Let X be a random variable with cdf F_X

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

no flat in (a, b)

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

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



Proof

Let $Z = F^{-1}(Y)$

$F(F^{-1}(t)) = t, 0 \le t \le 1$

$F^{-1}(F(s)) = s, a < s < b$

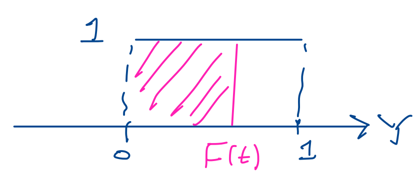
$$F_Z(t) = P(\underline{\underline{Z}} \le t)$$

$$= P(F_X^{-1}(Y) \le t)$$

$\because F$ is increasing

$$= P(\underbrace{F_X(F_X^{-1}(Y))}_{Y} \le F_X(t))$$

$$= P(\underbrace{\underline{\underline{Y}}}_{\sim \mathrm{Unif}(0,1)} \le \underbrace{F_X(t)}_{\in [0,1]})$$



$$= F_X(t)$$

$\Rightarrow Z = X$ in distribution

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

∎

Change of Variables

Example

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

Find the distribution of $Y = (1-X)^3$.

Exercise

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

Find the pdf of $Y = X^2$. $y > 0$ $\quad t \geq 0$ $\quad$ otherwise $F_Y = 0$.

$$F_Y(t) = \mathbb{P}(Y \leq t) = \mathbb{P}(X^2 \leq t)$$

$$= \mathbb{P}(-\sqrt{t} \leq X \leq \sqrt{t}) \leftarrow$$

$$= \mathbb{P}(X \leq \sqrt{t}) - \mathbb{P}(X < -\sqrt{t})$$

$$= \mathbb{P}(X \leq \sqrt{t}) = F_X(\sqrt{t})$$

$$f_Y(t) = \frac{d}{dt} F_Y(t) = \frac{d}{dt} F_X(\sqrt{t})$$

$$= F_X'(\sqrt{t}) \cdot (\sqrt{t})'$$

$0 < \sqrt{t} < 1$

$$= f_X(\sqrt{t}) \cdot \frac{1}{2\sqrt{t}} = \cancel{4}^{2} \cdot (\sqrt{t})^{\cancel{3}\,2} \cdot \frac{1}{\cancel{2}\cancel{\sqrt{t}}} = 2t .$$

0 $\quad$ otherwise .

$$f_Y(t) = \begin{cases} 2t , & 0 < t < 1 \\ 0 , & \text{o.w.} \end{cases}$$

$Y = u(X)$, u is strictly increasing.

$$F_Y(t) = \mathbb{P}(u(X) \leq t) \qquad (v = u^{-1})$$

$$= \mathbb{P}(X \leq v(t))$$

$$= \underbrace{F_X(v(t))}$$

$$f_Y(t) = f_X(v(t)) \cdot v'(t)$$

Section 2. Transformations of Two Random Variables

Transformations of Two Random Variables

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

$x_1 x_2$

ex) $\begin{cases} Y_1 = X_1 + X_2 = u_1(X_1, X_2) \\ Y_2 = X_1 \cdot X_2 = u_2(X_1, X_2) \end{cases}$

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

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

(Change of Variables)

where J is the Jacobian given by

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

determinant.

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 $Y_1 = \frac{X_1}{X_2}$ and $Y_2 = X_2$.

$$\begin{cases} Y_1 = u_1(X_1, X_2) = \dfrac{X_1}{X_2} \\ Y_2 = u_2(X_1, X_2) = X_2 \end{cases} \Rightarrow \quad X_1 = X_2 \cdot Y_1 = Y_1 \cdot Y_2$$

$$\Rightarrow \begin{cases} X_1 = Y_1 \cdot Y_2 = v_1(Y_1, Y_2) \\ X_2 = Y_2 = v_2(Y_1, Y_2) \end{cases}$$

$$|J| = \begin{vmatrix} \frac{\partial x_1}{\partial y_1} & \frac{\partial x_1}{\partial y_2} \\ \frac{\partial x_2}{\partial y_1} & \frac{\partial x_2}{\partial y_2} \end{vmatrix} = \begin{vmatrix} y_2 & y_1 \\ 0 & 1 \end{vmatrix} = |y_2|$$

$$f_{Y_1,Y_2}(y_1, y_2) = \underbrace{f_{X_1,X_2}(y_1 \cdot y_2, y_2)} \cdot |y_2| = \begin{cases} 2y_2, & 0 < y_1 y_2 < y_2 < 1 \\ 0, & \text{o.w.} \end{cases}$$

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.

joint

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

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 joint pdfs, then the joint pdf is the product. of marginals

If $X_1, X_2, \cdots, X_n$ are independent and have the same distribution, we say they are i.i.d. or a random sample of size n from that common distribution.

Independent, identically distributed

For $\{X_{i_1} \in A_{i_1}\}, \cdots, \{X_{i_k} \in A_{i_k}\}$

$$P(X_{i_1} \in A_{i_1}, \cdots, X_{i_k} \in A_{i_k}) = P(X_{i_1} \in A_{i_1}) \cdots P(X_{i_k} \in A_{i_k})$$

Independent random variables

$Y \sim Exp(\lambda)$ $\qquad$ $\mathbb{P}(Y \underset{(=)}{>} t) = e^{-\lambda t}$

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

Example

Let X_1, X_2, X_3 be a $\overset{\text{Indep.}}{}$ random sample from a distribution with pdf

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

Exp (1)

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

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

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

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

$= e^{-5} \cdot (1 - e^{-1})(1 - e^{-2})(1 - e^{-4})$.

$$\mathrm{Var}(X) = \mathbb{E}[(X - \mathbb{E}X)^2]$$

$$\mathrm{Cov}(X,Y) = \mathbb{E}[\underbrace{(X - \mathbb{E}X)}\cdot(\underbrace{Y - \mathbb{E}[Y]})]$$

$$= \mathbb{E}[\bar{X} \cdot \bar{Y}]$$

Expectation and Variance

Theorem

without indep.

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

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

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

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

$\bar{X} = X - \mathbb{E}X$, $\bar{Y} = Y - \mathbb{E}Y$

Note

$$\mathrm{Var}(X+Y) = \mathrm{Var}(\bar{X} + \bar{Y}) \qquad \mathbb{E}\bar{X} = \mathbb{E}\bar{Y} = 0$$

$$\mathbb{E}[(\overbrace{X+Y}^{\shortparallel} - \underbrace{\mathbb{E}[X+Y]})^2] \to \mathbb{E}[(\bar{X} + \bar{Y})^2]$$

$$= \mathbb{E}[\bar{X}^2 + 2\bar{X}\cdot\bar{Y} + \bar{Y}^2]$$

$$= \mathbb{E}[\bar{X}^2] + 2\,\mathbb{E}[\bar{X}\bar{Y}] + \mathbb{E}[\bar{Y}^2]$$

?

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

Exercise

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

Exercise 5.2.6

$X_1 \sim \text{Gamma}(\alpha, \theta)$ Indep.
$X_2 \sim \text{Gamma}(\beta, \theta)$

$$f_{X_1}(x_1) = \frac{1}{\theta^\alpha \Gamma(\alpha)} x_1^{\alpha-1} e^{-\frac{x_1}{\theta}}, \quad x_1 > 0$$

$$f_{X_2}(x_2) = \frac{1}{\theta^\beta \Gamma(\beta)} x_2^{\beta-1} e^{-\frac{x_2}{\theta}}$$

$$\frac{X_1}{X_1+X_2} < 1$$

$$W = \frac{X_1}{X_1+X_2}$$

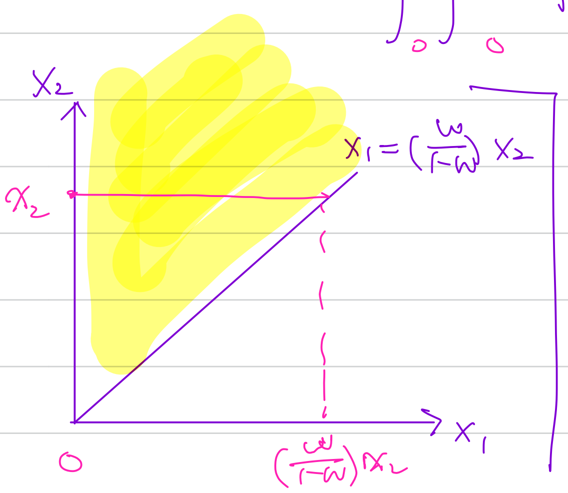
$1 > w > 0$

$$F_W(w) = P(W \leq w) = P\left(\frac{X_1}{X_1+X_2} \leq w\right)$$

$$= P(X_1 \leq w(X_1+X_2))$$

$$= P((1-w)X_1 \leq w X_2) = P\Big(\underbrace{X_1 \leq \Big(\overset{>0}{\frac{w}{1-w}}\Big) X_2}\Big)$$

$$= \int_0^\infty \int_0^{\left(\frac{w}{1-w}\right)x_2} f_{X_1}(x_1) \cdot f_{X_2}(x_2)\, dx_1\, dx_2$$



$$= \int_0^\infty \left(\int_0^{\left(\frac{w}{1-w}\right)x_2} f_{X_1}(x_1)\, dx_1 \right) \cdot f_{X_2}(x_2)\, dx_2$$

$$= \int_0^\infty F_{X_1}\left(\left(\frac{w}{1-w}\right)x_2\right) \cdot f_{X_2}(x_2)\, dx_2$$

$$f_W(w) = \int_0^\infty \frac{d}{dw} F_{X_1}\left(\left(\frac{w}{1-w}\right)x_2\right) \cdot f_{X_2}(x_2)\, dx_2$$

$$= \int_0^\infty f_{X_1}\left(\frac{w}{1-w} x_2\right) \cdot \frac{x_2}{(1-w)^2} f_{X_2}(x_2)\, dx_2$$

$$= \frac{1}{\theta^\alpha \Gamma(\alpha)} \frac{1}{\theta^\beta \Gamma(\beta)} \int_0^\infty \left(\left(\frac{w}{1-w}\right)x_2\right)^{\alpha-1} e^{-\frac{\left(\frac{w}{1-w}\right)x_2}{\theta}} \cdot \frac{x_2}{(1-w)^2} x_2^{\beta-1} e^{-\frac{x_2}{\theta}}\, dx_2$$

$$= \frac{1}{\theta^{\alpha+\beta}\Gamma(\alpha)\Gamma(\beta)} \frac{w^{\alpha-1}}{(1-w)^{\alpha+1}} \int_0^\infty x_2^{\alpha+\beta-1} e^{-\frac{x_2}{\theta}\cdot\left(\underbrace{\frac{w}{1-w}+1}_{=\frac{1}{1-w}}\right)} d x_2$$

$$\int_0^\infty x_2^{(\alpha+\beta)-1} \cdot e^{-\underbrace{\frac{x_2}{\theta(1-w)}}_{=t}} d x_2 \qquad d x_2 = \theta(1-w)\, dt$$

$$= \int_0^\infty \left(\theta(1-w)\, t\right)^{\alpha+\beta-1} e^{-t}\, \theta(1-w)\, dt$$

$$= \theta^{\alpha+\beta} \cdot (1-w)^{\alpha+\beta} \underbrace{\int_0^\infty t^{\alpha+\beta-1} e^{-t}\, dt}_{= \Gamma(\alpha+\beta)}$$

$$= \frac{1}{\cancel{\theta^{\alpha+\beta}}\Gamma(\alpha)\Gamma(\beta)} \frac{w^{\alpha-1}}{(1-w)^{\alpha+1-\beta}} \cdot \cancel{\theta^{\alpha+\beta}}\, \cancel{(1-w)^{\alpha+\beta}}\, \Gamma(\alpha+\beta)$$

$$= \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \cdot w^{\alpha-1} (1-w)^{\beta-1}$$

Beta dist.

$0 < w < 1$

$$\int_0^1 w^{\alpha-1} (1-w)^{\beta-1}\, dw = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} =: B(\alpha, \beta)$$

Beta function

Section 4.
The Moment-Generating Function Technique

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

Fact X, Y, $M_X(t) = M_Y(t)$ for $-\delta < t < \delta$ for some $\delta > 0$

$\Rightarrow$ $F_X(t) = F_Y(t)$ $\forall t \in \mathbb{R}$

$\Rightarrow$ X, Y have the same distribution

The Moment-Generating Function

i.i.d = Independent and identically distributed

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**~~ Corollary

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

$a_1 = \cdots = a_n = \frac{1}{n}$

Proof

$$
\begin{aligned}
M_Y(t) &= \mathbb{E}[e^{tY}] \\
&= \mathbb{E}[e^{t(a_1X_1 + \cdots + a_nX_n)}] \\
&= \mathbb{E}[e^{ta_1X_1 + ta_2X_2 + \cdots + ta_nX_n}] \\
&= \mathbb{E}[e^{(ta_1)\cdot X_1} \cdot e^{(ta_2)X_2} \cdot \cdots \cdot e^{(t\cdot a_n)X_n}] \\
&= \mathbb{E}[e^{(ta_1)X_1}] \cdot \cdots \cdot \mathbb{E}[e^{(ta_n)X_n}] \\
&= M_{X_1}(ta_1) \cdots M_{X_n}(ta_n)
\end{aligned}
$$

$\square$

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.

$X \sim Ber(p)$

① $M_X(t) = \mathbb{E}[e^{tX}] = \sum_x e^{tx} \cdot p(x) = e^{t\cdot 0} \cdot (1-p) + e^{t \cdot 1} \cdot p$

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

② $M_Y(t) = (M_X(t))^n = ((1-p) + p \cdot e^t)^n$

③ $W \sim Bin(n, p)$

$M_W(t) = \mathbb{E}[e^{tW}] = \sum_x e^{tx} \cdot p(x)$

$= \sum_{x=0}^{n} e^{tx} \cdot \binom{n}{x} \cdot p^x (1-p)^{n-x}$

$= \sum_{x=0}^{n} \binom{n}{x} (e^t \cdot p)^x (1-p)^{n-x} = (e^t p + (1-p))^n$

Binomial

$Y \sim Bin(n, p)$

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.

① $X \sim$ Exp with θ

$$M_X(t) = \mathbb{E}[e^{tX}] = \int_0^\infty e^{tx} \frac{1}{\theta} e^{-\frac{x}{\theta}}\, dx$$

$$= \frac{1}{\theta} \int_0^\infty e^{-(\frac{1}{\theta} - t)x}\, dx \quad \leftarrow \quad \frac{1}{\theta} - t > 0$$

$t < \frac{1}{\theta}$

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

② $M_Y(t) = (M_X(t))^n = (1 - \theta t)^{-n} \leftarrow$ Gamma ✓ Exer.

- $M_X(t) = M_Y(t)$ for $t \in (-\delta, \delta)$, $\delta > 0$

  if and only if $X \overset{d}{=} Y$ ($F_X(t) = F_Y(t)$ $\forall t$)

- $X_1, X_2, \ldots, X_n$ i.i.d RVs (random sample of size n)

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

  $\bar{X} = \frac{1}{n}(X_1 + \cdots + X_n) \Rightarrow M_{\bar{X}}(t) = (M_X(\frac{t}{n}))^n$

  Sample mean

Exercise

$X \sim \text{Poisson}(\lambda)$ $\quad f_X(k) = e^{-\lambda} \frac{\lambda^k}{k!}$ $\quad k = 0, 1, 2, \ldots$

$$M_X(t) = \mathbb{E}[e^{tX}] = \sum_{k=0}^{\infty} e^{tk} \cdot e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \underbrace{\sum_{k=0}^{\infty} \frac{(e^t \lambda)^k}{k!}}_{e^{e^t \lambda}}$$

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

Let X_1, X_2, X_3 be independent Poisson with means 2, 1, 4.

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

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

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

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

$\Rightarrow Y \sim \text{Poisson}(7)$

$\{X_i\}$ i.i.d	$\sum_{i=1}^{n} X_i$
Ber (p)	Bin (n, p)
Bin (m, p)	Bin $(n \cdot m, p)$
Pois (λ)	Pois (λn)
Geom (p)	NegBin (n, p)
Exp (λ)	Gamma (n, λ)

- $Y \sim \text{Bin}(n, p)$

  $\Rightarrow Y \overset{d}{=} X_1 + X_2 + \cdots + X_n$

  $\{X_i\}$ i.i.d. $\text{Ber}(p)$

- $Y \sim \text{Pois}(100)$

  $Y \overset{d}{=} X_1 + \cdots + X_{100}$

  $\{X_i\}$ i.i.d. $\text{Pois}(1)$

$N(\mu, \sigma^2)$ $\quad$ $Normal(n \cdot \mu, n\sigma^2)$

Section 6.
The Central Limit Theorem

$$\mathbb{E}[\overline{X}] = \mathbb{E}\left[\frac{1}{n}(X_1 + X_2 + \cdots + X_n)\right] = \frac{1}{n}\underline{\mathbb{E}[X_1 + \cdots + X_n]}$$

$$= \frac{1}{n}\cdot\left(\mathbb{E}X_1 + \mathbb{E}X_2 + \cdots + \mathbb{E}X_n\right) = \frac{1}{\cancel{n}}\cdot\cancel{n}\cdot\mathbb{E}[X_1] = \mu$$

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

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

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

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

$$W = \frac{\overline{X} - \mathbb{E}[\overline{X}]}{\sqrt{\mathrm{Var}(\overline{X})}} = \frac{\overline{X} - \mu}{\sqrt{\sigma^2/n}} = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

normalization

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

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

$$\mathbb{E}[W] = \mathbb{E}\left[\frac{1}{\sigma/\sqrt{n}}(\overline{X} - \mu)\right] = \frac{1}{\sigma/\sqrt{n}}\cdot\left(\mathbb{E}[\overline{X}] - \mu\right) = 0$$

$$\mathrm{Var}(W) = \mathrm{Var}\left(\frac{1}{\sigma/\sqrt{n}}(\overline{X} - \mu)\right) = \frac{1}{\sigma^2/n}\cdot\underbrace{\mathrm{Var}(\overline{X} - \mu)}_{=\ \mathrm{Var}(\overline{X}) = \frac{\sigma^2}{n}} = 1.$$

$X_1, X_2, X_3, \ldots$: i.i.d $\mu < \infty$, $\sigma^2 < \infty$

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

$W_n = W = \frac{\bar{X}_n - \mu}{\sigma/\sqrt{n}}$ $W_n \Rightarrow N(0,1)$ in distribution as $n \to \infty$

The Central Limit Theorem

Gaussian universality

Normal Dist = Gaussian

Theorem

If μ and σ^2 are finite, then the distribution of W_n converges to that of the standard normal distribution as $n \to \infty$.

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

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

Idea

$\mu = 0$, $\sigma^2 = 1$ for simplicity.

$$W = \frac{\bar{X} - 0}{1/\sqrt{n}} = \sqrt{n} \cdot \bar{X} = \frac{X_1 + \cdots + \cancel{X_n}}{\sqrt{n}}$$

$$M_W(t) = \left(M_X\left(\frac{t}{\sqrt{n}}\right) \right)^n$$

① $M_X(0) = \mathbb{E}[e^{0 \cdot X}] = 1$

$M_X'(0) = 0$

$M_X''(0) = \mathbb{E}[X^2] - (\overset{0}{\mathbb{E}X})^2 = \sigma^2 = 1.$

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

$$\approx \underbrace{1 + \frac{1}{2}t^2}$$

② $$M_W(t) \approx \left(1 + \frac{1}{2}\cdot\left(\frac{t}{\sqrt{n}}\right)^2\right)^n = \left(1 + \frac{t^2}{2n}\right)^n$$

$$\longrightarrow e^{t^2/2} = M_Z(t)$$

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

The Central Limit Theorem

(CLT)

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

$$W = \frac{\overline{X} - \mu}{\sigma/\sqrt{n}} = \frac{\overline{X} - 15}{2/5} \Rightarrow N(0,1)$$

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

$$= \mathbb{P}\left(-\frac{3/5}{2/5} < \frac{\overline{X} - 15}{2/5} < \frac{3/5}{2/5}\right)$$

$$= \mathbb{P}\left(-\frac{3}{2} < W < \frac{3}{2}\right) = \mathbb{P}(W < 1.5) - \mathbb{P}(W \leq -1.5)$$

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

$$= 2\cdot\underline{\Phi(1.5)} - 1$$

using Table

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

$$\mathbb{E}[X] = \int_0^2 \underbrace{x \cdot \frac{x^3}{4}} \, dx = \left[\frac{1}{20} x^5\right]_0^2$$

$$= \frac{32}{20} = \frac{8}{5} = 1.6 = \mu$$

$$\mathrm{Var}(X) = \frac{8}{3} - \left(\frac{8}{5}\right)^2 = \frac{200-192}{75} = \frac{8}{75} = \sigma^2$$

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

$$\frac{\sqrt{\frac{8}{75}}}{\sqrt{25}} = \frac{2\sqrt{2}}{\sqrt{3}\cdot 5 \cdot 5}$$

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

$$= \mathbb{P}\left(\frac{-0.1}{} \leq \frac{\overline{X} - 1.6}{\sqrt{\frac{8}{75}} / \sqrt{25}} \leq \frac{0.05}{}\right) \approx \Phi(b) - \Phi(a)$$

$$\text{"} = a \qquad\qquad \text{"} = b$$

$$P\left(\frac{2.5-5}{\sqrt{10\cdot\frac{1}{2}\cdot\frac{1}{2}}}\le Z\le\frac{5.5-5}{\sqrt{10\cdot\frac{1}{2}\cdot\frac{1}{2}}}\right)\approx P(2.5<Y<5.5)=P(Y=3,4,5)$$

$\parallel$

<u>Example</u> $Y\sim Bin(10,\frac{1}{2})$ $\quad P(3\le Y<6)\approx ?$

Approx. by CLT?

In general, $Y\sim Bin(n,p)$ n large.

$Y=X_1+\cdots+X_n$, $\{X_i\}$ i.i.d. $Ber(p)$

$\bar{X}=\frac{Y}{n}$, $\mu=p$, $\sigma^2=p(1-p)$

$$P(W\le x)\approx P(Z\le x)$$

$$W=\frac{\bar{X}-\mu}{\sigma/\sqrt{n}}=\frac{Y/n-p}{\sqrt{p(1-p)}/\sqrt{n}}=\frac{Y-np}{\sqrt{np(1-p)}}$$

$\Rightarrow N(0,1)$

$$P(Y\le k)$$

$$=P\left(\frac{Y-np}{\sqrt{np(1-p)}}\le\frac{k-np}{\sqrt{np(1-p)}}\right)\approx P\left(Z\le\frac{k-np}{\sqrt{np(1-p)}}\right)$$

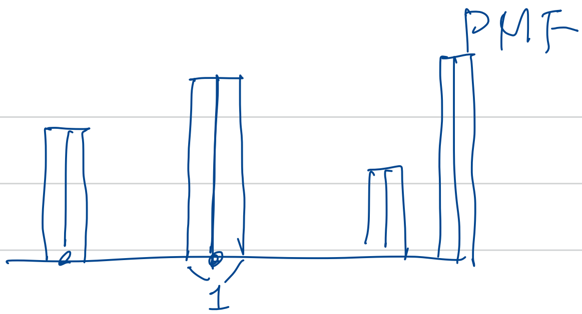
$Y\sim Bin(100,\frac{1}{2})$

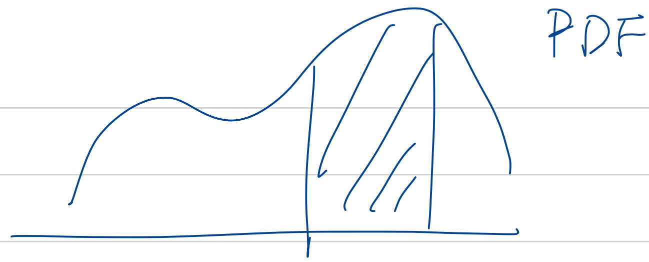
$$P(Y=53)=P\left(\frac{Y-50}{\underbrace{\sqrt{100\cdot\frac{1}{2}\cdot\frac{1}{2}}}_{=5}}=\frac{53-50}{5}\right)$$

$$\approx P\left(Z=\frac{3}{5}\right)=0$$

$$P(Y=53)=P(52.5<Y<53.5)$$ half unit correction.

$$\approx P\left(\frac{52.5-50}{5}\le Z<\frac{53.5-50}{5}\right)$$





$Y \sim \text{Poisson}(100)$ $\quad Y = X_1 + \cdots + X_{100}$,

$\{X_i\}$ i.i.d $\quad \text{Pois}(1)$

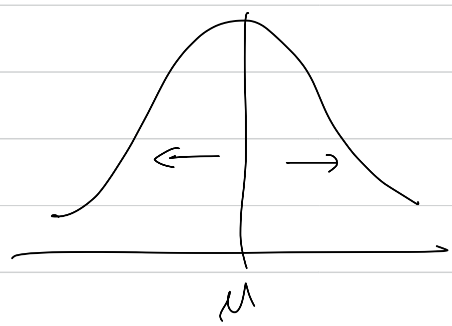
$$W = \frac{\bar{X} - 1}{1/\sqrt{100}} = \sqrt{100}\left(\frac{Y}{100} - 1\right) = \frac{Y - 100}{\sqrt{100}} = \frac{Y - 100}{10}$$

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

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

$\underline{\text{CLT}} \Leftarrow$ Typical Fluctuation.

$\bar{X} \approx \mu$



Is $\bar{X}$ close μ?

The typical behavior of $\bar{X}$ is μ?

: Law of Large Numbers.

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.

Section 8.
Chebyshev's Inequality and Convergence in Probability

Simple case $\quad X$: a RV $\quad X \geq 0$

$x \geq k$

$\frac{x}{k} \geq 1$

$$\mathbb{P}(X \geq k) = \sum_{x:\, x \geq k} 1 \cdot f_X(x)$$

$$\leq \sum_{x:\, x\geq k} \frac{x}{k} f_X(x) = \frac{1}{k} \sum_{x:\, x \geq k} x \cdot f_X(x)$$

$$\leq \frac{1}{k} \underbrace{\boxed{\sum_x x \cdot f_X(x)}}_{= \mathbb{E}[X]} = \frac{\mathbb{E}X}{k}$$

Chebyshev's Inequality

Replace X with $|X-\mu|^2$

Theorem

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

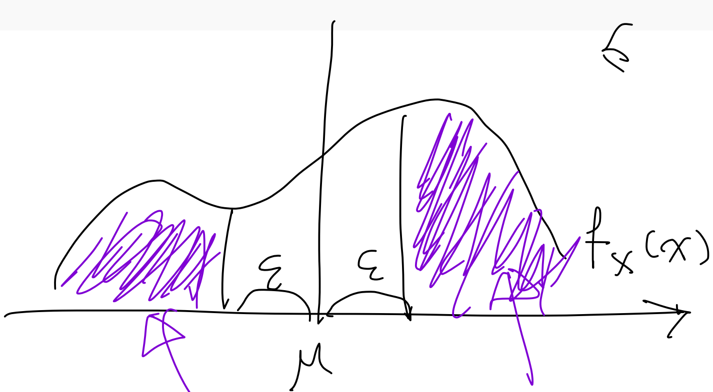
$$\mathbb{P}(|X-\mu| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}.$$

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

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

$$\mathbb{P}(|X-\mu| \geq \varepsilon) = \mathbb{P}(|X-\mu|^2 \geq \varepsilon^2)$$

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



$\mathbb{P}(|X-\mu| \geq \varepsilon)$

$|X-\mu| \geq \varepsilon$

$\Leftrightarrow \quad X-\mu \geq \varepsilon$, or

$X - \mu \leq -\varepsilon$,

$X \geq \mu + \varepsilon$ or $X \leq \mu - \varepsilon$

Suppose X is a nonnegative RV, $\mathbb{P}(X \geq 0) = 1$.
For any $k > 0$,

$$\mathbb{P}(X \geq k) \leq \frac{\mathbb{E}[X]}{k}$$

Chebyshev's Inequality.

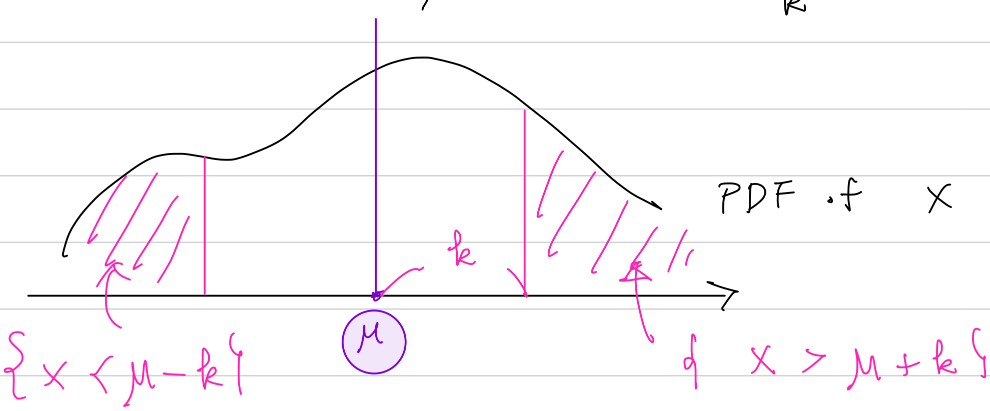
Application

X : any RV with finite variance.
$\mathbb{E}[X] = \mu$, $Var(X) = \sigma^2 < \infty$.
$Y = |X - \mu| \geq 0$

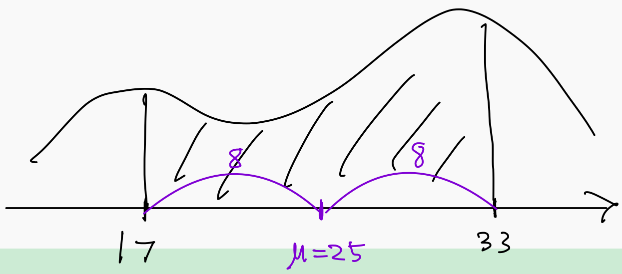
$$\begin{aligned}
\mathbb{P}(|X-\mu| \geq k) = \mathbb{P}(Y \geq k) &= \mathbb{P}(Y^2 \geq k^2) \\
&= \mathbb{P}(|X-\mu|^2 \geq k^2) \leq \frac{\mathbb{E}[|X-\mu|^2]}{k^2} \\
&= \frac{Var(X)}{k^2} = \frac{\sigma^2}{k^2}
\end{aligned}$$

Tail Probability.

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



Chebyshev's Inequality



Example

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

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

$$\mathbb{P}(17 < X < 33) = \mathbb{P}(|X - \overset{25}{\mu}| < 8)$$

$$= 1 - \mathbb{P}(|X-\mu| \geq \underline{8})$$

$$\geq 1 - \frac{\mathrm{Var}(X)}{8^2} = 1 - \frac{16}{64} = \frac{3}{4}.$$

$X_1, X_2, \cdots, X_n$ i.i.d. with $\mathbb{E}[X_1] = \mu < \infty$, $Var(X_1) = \sigma^2 < \infty$

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

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

$Var(\overline{X}) = \frac{1}{n}\cdot\sigma^2$

$$\mathbb{P}(|\overline{X} - \mu| \geq k) \leq \frac{Var(\overline{X})}{k^2} = \frac{\sigma^2}{n \cdot k^2}$$

$\forall$ all $k > 0$

The Law of Large Numbers

$$\lim_{n\to\infty} \mathbb{P}(|\overline{X}_n - \mu| \geq k) = 0 \quad \forall k > 0.$$

Law of Large Numbers

" $\overline{X}$ converges to μ as $n \to \infty$ " in probability.

Definition

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

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

Recall (CLT)

$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$ converges to $N(0,1)$ as $n \to \infty$

in distribution

That is $\mathbb{P}\left(\frac{\overline{X} - \mu}{\sigma/\sqrt{n}} \leq x\right) \longrightarrow \mathbb{P}(Z \leq x)$

The Law of Large Numbers

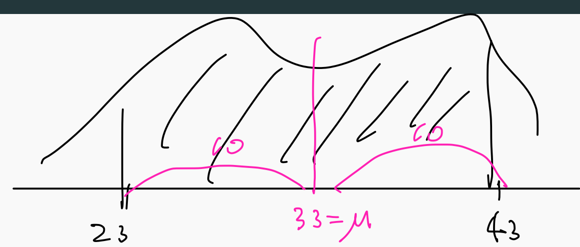
Theorem

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

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

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

Exercise



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

1. A lower bound for $\mathbb{P}(23 < X < 43)$. $= \mathbb{P}(|X - \overset{33}{\mu}| < 10) = 1 - \mathbb{P}(|X-\mu| \geq 10)$

$$\geq 1 - \frac{Var(X)}{10^2}$$

$$= 1 - \frac{16}{100} = \frac{84}{100}$$

2. An upper bound for $\mathbb{P}(\underbrace{|X - 33|}_{33} \geq 14)$.

$$\leq \frac{Var(X)}{14^2} = \frac{16}{196}.$$