MS Theory - I

Lecture - 3

Review :

§ 1.3 Conditional Probability and Independence

Definition 1.3.2 If A and B are events in S, and $P(B) > 0$, then the *conditional probability of A given B*, written $P(A|B)$, is

$$P(A|B) = \frac{P(A \cap B)}{P(B)}. \tag{1.3.1}$$

RIP Buttercup

Definition 1.3.7 Two events, A and B, are *statistically independent* if

$$P(A \cap B) = P(A)P(B). \tag{1.3.8}$$

Also: $P(A|B) = P(A)$ $\quad P(B) > 0 \not\Rightarrow$ $\quad A \perp B$

$P(B|A) = P(B)$ $\quad P(A) > 0 \not\Rightarrow$ $\quad A \perp B$

Theorem 1.3.9 *If A and B are independent events, then the following pairs are also independent:*

a. A *and* B^c,
b. A^c *and* B,
c. A^c *and* B^c.

Definition 1.3.12 A collection of events $A_1, \ldots, A_n$ are *mutually independent* if for any subcollection $A_{i_1}, \ldots, A_{i_k}$, we have

$$P\left(\bigcap_{j=1}^{k} A_{i_j}\right) = \prod_{j=1}^{k} P\left(A_{i_j}\right).$$

To show Events A, B and C are mutually independent

$$\left.\begin{aligned} P(ABC) &= P(A \cap B \cap C) = P(A)P(B)P(C) \\ P(AB) &= P(A \cap B) = P(A)P(B) \\ P(AC) &= P(A \cap C) = P(A)P(C) \\ P(BC) &= P(B \cap C) = P(B)P(C) \end{aligned}\right] \Longrightarrow \text{A, B, C mutually independent}$$

§1.4 Random Variables

Definition 1.4.1 A *random variable* is a function from a sample space S into the real numbers.

Notation: random variables denoted with uppercase letters; realized value of variable (range) denoted by lower case.

Random Variable X takes on values x.

Flip coin 3 times

$S =$

HHH	HHT	HTH	THH
HTT	THT	TTH	TTT

Define random variable X = number heads in 3 tosses

S	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
$X(s)$	3	2	2	2	1	1	1	0

$P(X=3) = 1/8$, $P(X=2) = 3/8$, $P(X=1) = 3/8$, $P(X=0) = 1/8$

Range for $X = \mathcal{X} = \{0,1,2,3\}$

Assuming a fair coin:

x	0	1	2	3
$P_X(X=x)$	$1/8$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

Random Variable; realized value: 0,1,2 or 3

$P(X=3)$ — Random Variable; realized value

Examples

(Random Variable maps $S \to \mathbb{R}^1$)

$\mathbb{R}^1 =$ ⟵|||||||0|||||||⟶ = real numbers

Experiment	Random Variable, X	Range of Random Variable (Sample Space of X)
Give new treatment to patients with common cold.	X = number of 'cures' in n patients	$0, 1, 2, \ldots, n$
Give a patient with high blood pressure new medication	X = change in systolic blood pressure	$-\infty < X < \infty$
Toss 2 die	X = sum of numbers	$2, 3, \ldots 12$

§ 1.5 Distribution Functions

Cumulative distribution function (cdf)

Definition 1.5.1 The *cumulative distribution function* or *cdf* of a random variable X, denoted by $F_X(x)$, is defined by

$$F_X(x) = P_X(X \le x), \quad \text{for all } x.$$

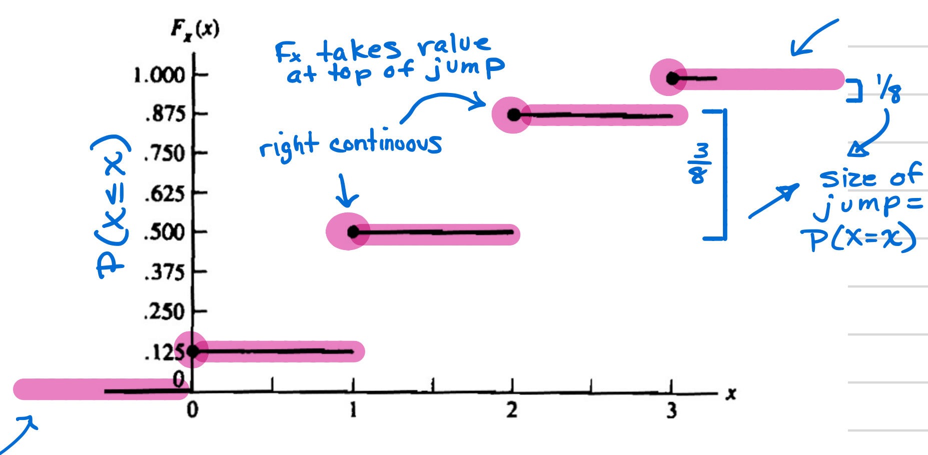
Note: for all x, not just those in sample space $-\infty < X < \infty$.

Example: Toss coin 3 times

X = # heads $\quad S = \{0, 1, 2, 3\}$

$$F_X(x) = P_X(X \le x) = P(X \le x) = \begin{cases} 0 & \text{if } \underline{-\infty} < x < 0 \\ 1/8 & \text{if } 0 \le x < 1 \\ 1/2 & \text{if } 1 \le x < 2 \\ 7/8 & \text{if } 2 \le x < 3 \\ 1 & \text{if } 3 \le x < \underline{\infty} \end{cases}$$

PROBABILITY THEORY



defined for all values of x, not just

Figure 1.5.1. *Cdf of Example 1.5.2*

$P(X \le 3) = 1 \qquad P(X \le 2) = 7/8 \qquad P(X < 2) = 1/2$

$P(X \le 25.2) = 1 \qquad P(X \le 2.6) = 7/8 \qquad P(X < 1.2) = 1/2$

Note:
- cdf, F_X, defined for all x $(-\infty < x < \infty)$, not just those in Sample Space.
- $P(X \le x)$ ⟹ defined (equal sign) to be right continuous
  ↑ equal sign

  right continuous: ft'n is continuous when approached from right

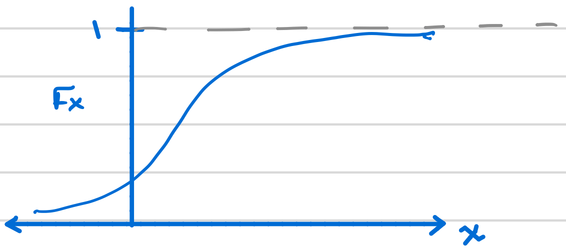
  right continuous: F_X takes value at top of jump
- Size of jump at any point $X = x$ is equal to $P(X = x)$
- as $x \to -\infty$ F_X goes to 0.
- as $x \to \infty$ F_X goes to 1.
- $F_X = P(X \le x)$ is nondecreasing ft'n.

Theorem 1.5.3 *The function $F(x)$ is a cdf if and only if the following three conditions hold:*

a. $\lim_{x\to-\infty} F(x) = 0$ *and* $\lim_{x\to\infty} F(x) = 1$.

b. *$F(x)$ is a nondecreasing function of x.*

c. *$F(x)$ is right-continuous; that is, for every number x_0, $\lim_{x\downarrow x_0} F(x) = F(x_0)$.*



cdfs may also be continuous (more soon).

Example: Clinical trial enrollment example:

- Enroll 50 individuals for a PK (pharmacokinetics) trial to measure drug levels for an antiretroviral drug to treat patients with HIV disease.
- The trial is balanced by whether subjects express a CYP3A5 gene and race
- Let p = probability of gene expression
- Define a random variable X = number of genetic screens required to find an expressor.
- Then for any $x = 1, 2, \ldots$

$$P(X=x) = \underbrace{(1-p)^{x-1}}_{(x-1)\text{ failures}} \underbrace{p}_{\text{followed by success}}$$

- Assume all participants are unrelated (independent)

$$P(X=1) = p$$
$$P(X=2) = (1-p)p$$
$$P(X=3) = (1-p)^2 p$$
$$P(X=4) = (1-p)^3 p$$
$$\vdots$$

$$P(X\le 3) = p + (1-p)p + (1-p)^2 p$$

$$P(X\le x) = \sum_{i=1}^{x} (1-p)^{i-1} p = p\sum_{i=1}^{x}(1-p)^{i-1}$$

Recall sum of geometric series $\sum_{k=1}^{n} t^{k-1} = \frac{1-t^n}{1-t}$ $t \neq 1$ (see lecture appendix for proof by induction)

$$F_X(x) = P(X\le x) = p\sum_{i=1}^{x}(1-p)^{i-1} = p\left(\frac{1-(1-p)^x}{(1-(1-p))}\right) = \cancel{p}\left(\frac{1-(1-p)^x}{\cancel{p}}\right)$$

$$= 1-(1-p)^x \qquad x = 1, 2, 3, \ldots$$

Show $F_X(x)$ is a cdf for $0<p<1$

a) $\lim_{x\to-\infty} F_X(x) = 0$ $\lim_{x\to\infty} F_X(x) = 1$

b) $F_X(x)$ is nondecreasing

c) right continuous; for every number x_0, $\lim_{x\downarrow x_0} F(x) = F(x_0)$

x approaches from right

Clinical trial enrollment continued:

$F_X(x) = P(X \le x) = 1 - (1-p)^x \quad x = 1, 2, \ldots \quad 0 < p < 1$

Show cdf: • $\lim_{x \to -\infty} 1 - (1-p)^x$ assuming possible x with non-zero probability are $x = 1, 2, 3, \ldots$ and $0 < p < 1$

since $P(X = x) = 0$ for any $x < 1$; $F_X(x) = 0 \ \forall x < 0$.
for all
the $\lim_{x \to -\infty} F(x) = 0$

• $\lim_{x \to \infty} 1 - (1-p)^x$ as $x \to \infty$ (as integer) $(1-p)^x \to 0$, since $0 < (1-p) < 1$

$\therefore \lim_{x \to \infty} 1 - (1-p)^x = 1$
therefore

• Show $F_X(x)$ is an increasing function in x
$F_X(x) = 1 - (1-p)^x$

- C&B note that with each increase in x we add another positive term, so
$p \sum_{i=1}^{x} (1-p)^{i-1}$ increases in x.

- we could also show the 1^{st} derivative is > 0 (see appendix) for $x = 1, 2, \ldots, 0 < p < 1$

• Show right continuos
$\lim_{x \downarrow x_0} F(x) = x_0$ for any number x_0

Since $x = 1, 2, 3, \ldots$ $F(x + \varepsilon) = F(x)$
for $\varepsilon > 0$, ε small

$\therefore \lim_{\varepsilon \downarrow 0} F_X(x + \varepsilon) = F_X(x) \Rightarrow$ right continuous.

$\therefore F_X$ is cdf

• Geometric Dist'n: (see page 621, 627)

Geometric(p) C&B p. 621

pmf	$P(X = x \mid p) = p(1-p)^{x-1}; \quad x = 1, 2, \ldots; \quad 0 \le p \le 1$
mean and variance	$EX = \frac{1}{p}, \quad \text{Var}\, X = \frac{1-p}{p^2}$

cdf of geometric $(p = 0.3)$ dist'n.

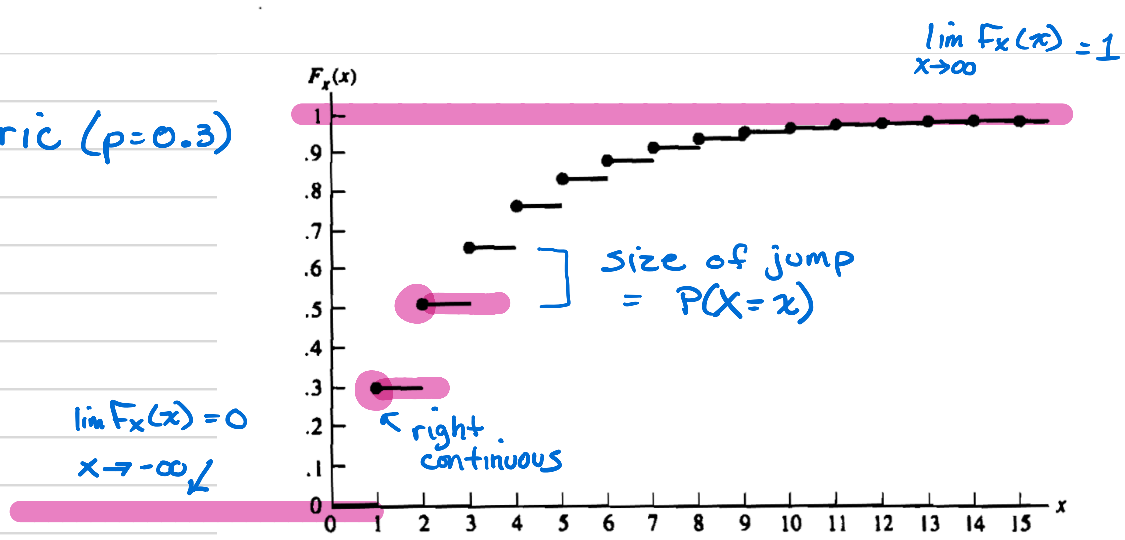


Figure 1.5.2. *Geometric cdf*, $p = .3$

Note cdf defined for $-\infty < x < \infty$

Example: Continuous cdf

$$F_X(x) = \frac{1}{1+e^{-x}}$$

Show limits of F_X

$$\lim_{x \to -\infty} \frac{1}{1+e^{-x}}$$

Note $\lim_{x \to -\infty} e^{-x} = \underbrace{\infty}_{\text{sloppy notation}}$

$$\therefore \lim_{x \to -\infty} \frac{1}{1+e^{-x}} = 0$$

$$\lim_{x \to \infty} \frac{1}{1+e^{-x}}$$

$\lim_{x \to \infty} e^{-x} = 0 \quad \therefore \lim_{x \to \infty} \frac{1}{1+e^{-x}} = 1$

Show nondecreasing: $\frac{d}{dx}\left(\frac{1}{1+e^{-x}}\right) = \frac{e^{-x}}{(1+e^{-x})^2} > 0$ (details appendix)

$\therefore F_X$ is nondecreasing

Show right continuous;

$F_X(x)$ is continuous in x, so therefore right continuous.

Special Case of logistic dist'n. $(\mu = 0, \beta = 1)$

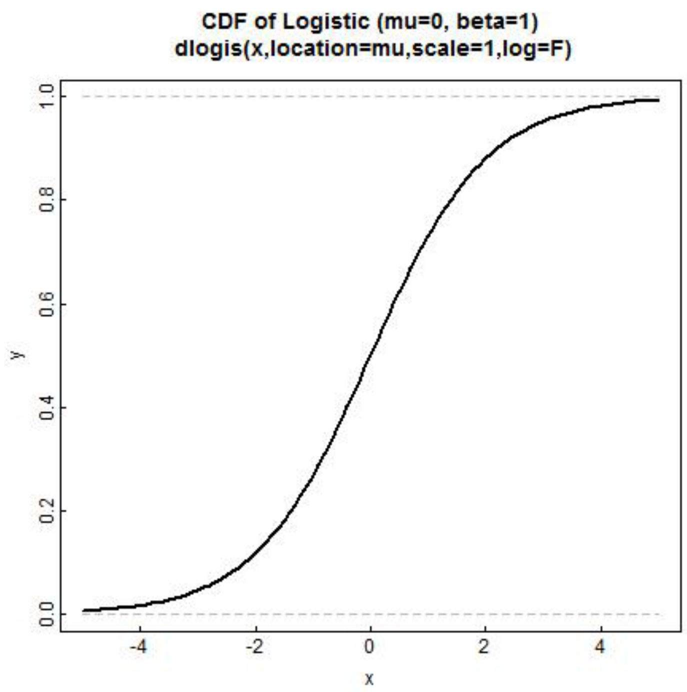
Logistic(μ, β)

pdf $\quad f(x|\mu,\beta) = \frac{1}{\beta}\frac{e^{-(x-\mu)/\beta}}{[1+e^{-(x-\mu)/\beta}]^2}, \quad -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad \beta > 0$

mean and variance $\quad \mathrm{E}X = \mu, \quad \mathrm{Var}\, X = \frac{\pi^2\beta^2}{3}$

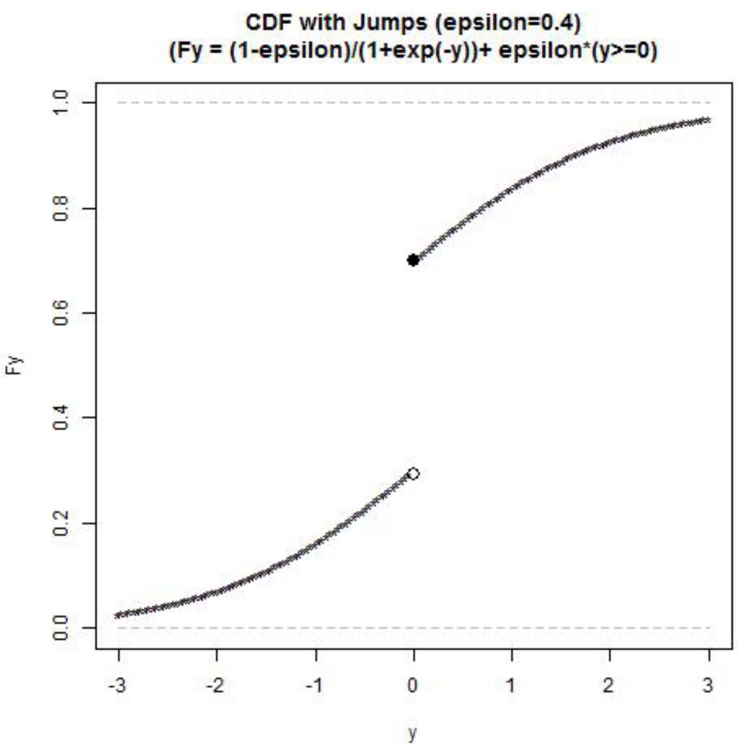
mgf $\quad M_X(t) = e^{\mu t}\Gamma(1-\beta t)\Gamma(1+\beta t), \quad |t| < \frac{1}{\beta}$

notes $\quad$ The cdf is given by $F(x|\mu,\beta) = \frac{1}{1+e^{-(x-\mu)/\beta}}$.



Example 1.5.6 cdf with jumps (sticking gauge).

$$F_Y(y) = \begin{cases} \frac{1-\varepsilon}{1+e^{-y}} & y < 0 \\ \varepsilon + \frac{(1-\varepsilon)}{1+e^{-y}} & y \geq 0 \end{cases}$$



Continuous or Discrete Random Variable, X.

Definition 1.5.7 A random variable X is *continuous* if $F_X(x)$ is a continuous function of x. A random variable X is *discrete* if $F_X(x)$ is a step function of x.

$F_X(x)$ is continuous ft'n of $x \not\Rightarrow$ Random Variable X is continuous.

$F_X(x)$ is step ft'n of $x \not\Rightarrow$ Random Variable X is discrete.

Further define identically distributed Random Variables

Definition 1.5.8 The random variables X and Y are *identically distributed* if, for every set $A \in \mathcal{B}^1$, $P(X \in A) = P(Y \in A)$.

well behaved σ-algebra
(events: intervals, countable unions, intersections of intervals, ...
→ Not "pathological cases."

Note: Two R.V.'s that are identically dist'd may not have $X = Y$.

example: flip a fair coin 3 times ($P(\text{head}) = P(\text{tails}) = \frac{1}{2}$)

X = # heads $\quad Y$ = # tails
(example 1.4.3)

Define random variable X = number heads in 3 tosses

s	HHH	HHT	HTH	THH	HTT	THT	TTH	TTT
X(s)	3	2	2	2	1	1	1	0
Y(s)	0	1	1	1	2	2	2	3

$X(s) \neq Y(s)$ for any sample points.

Assuming a fair coin:

x	0	1	2	3
$P_X(X=x)$	1/8	3/8	3/8	1/8
$P_Y(Y=y)$	1/8	3/8	3/8	1/8

§1.6 Density and Mass Functions

cdf = cumulative dist'n ft'n
$F_X(x) = P_X(X \le x) \quad \forall x$

denote cdfs by uppercase letters
$F(x)$ or $F_X(x)$

Also define
pmf = probability mass function (discrete R.V. X)
and
pdf = probability dist'n function (continuous R.V. X)

denote by lowercase letters
$f_X(x)$ or $f(x)$

Discrete R.V.s

Definition 1.6.1 The *probability mass function* (*pmf*) of a discrete random variable X is given by

$$f_X(x) = P(X = x) \quad \text{for all } x.$$

Example (Geometric Dist'n): (# trials to 1st success)

$$F_X(x) = \sum_{i=1}^{x} p(1-p)^{i-1} = 1-(1-p)^x$$

$$f_X(x) = P(X=x) = \underbrace{p}_{\text{prob 1 success}}\underbrace{(1-p)^{x-1}}_{\text{prob. } x-1 \text{ failures before 1st success}} \quad x=1,2,3,\dots$$

$$f_X(x) = P(X=x) = \begin{cases} (1-p)^{x-1}p & \text{for } x=1,2,\dots \\ 0 & \text{otherwise} \end{cases}$$

when we specify a pdf, we always note sample space (possible values of x).

if we define an indicator function:

$$I_{[1,2,\dots)}(x) = \begin{cases} 1 & \text{if } x \in [1,2,3,\dots] \\ 0 & \text{else} \end{cases}$$

$$f_X(x) = P(X=x) = (1-p)^{x-1} \cdot p \cdot I_{[1,2,3\dots]}(x)$$

using an indicator ft'n $I_{(\ )}(x)$ will make things easier soon...

cdf of geometric ($p=0.3$) dist'n.

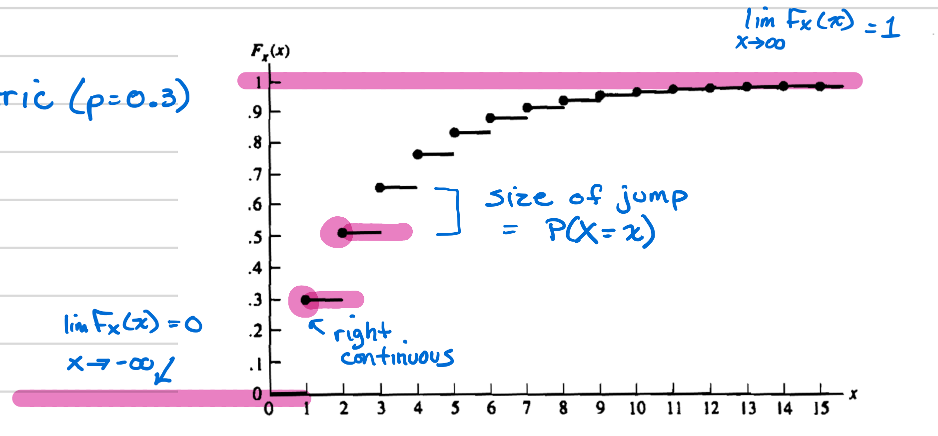


Figure 1.5.2. *Geometric cdf*, $p = .3$

Note cdf defined for $-\infty < x < \infty$

Remeber $P(X=x)$ is size of jump of cdf at x $(= f_X(x))$.

so $P(3 \le X \le 5)$
$= (1-p)^2p + (1-p)^3p + (1-p)^4p$
$= P(X=3) + P(X=4) + P(X=5)$
$= \sum_{k=3}^{5} f_X(k) = \sum_{k=3}^{5} (1-p)^{k-1}p$

For any a, b integers $\quad P(a \le X \le b) = \sum_{k=a}^{b} (1-p)^{k-1}p = \sum_{k=a}^{b} f_X(k)$

$F_X(b)$: and recall $F_X(b) = \text{cdf}$ $= \sum_{k=1}^{b} f_X(k) = \sum_{k=1}^{b} p(1-p)^{k-1}$

Discrete case $f_X(x) = P(X=x)$

pdf lower case

$F_X(x) = P(X \le x) = \sum_{k=1}^{x} f_X(k)$

cdf uppercase $P(a \le X \le b) = \sum_{k=a}^{b} f_X(k)$

Assuming (like geometric) a, b integers
$I_{[1,2,3,\dots)}(x) = 1$

Continuous R.Vs

Hmmm: If X is continuous $P(X=x)=0$

That is

$$P(x-\varepsilon < X \le x) = F_X(x) - F_X(x-\varepsilon) \quad \text{for } \varepsilon > 0$$

$$\therefore \; 0 \le P(X=x) \le \lim_{\varepsilon \downarrow 0} \left[F_X(x) - F_X(x-\varepsilon)\right] = 0 \quad (\text{since } F_X \text{ is continuous})$$

Since continuous use integrals instead of sums:

$$P(X \le x) = F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$$

cdf upper case   pdf lowercase

$$\therefore \; \frac{d}{dx} F_X(x) = f_X(x)$$

Definition 1.6.3 The *probability density function* or *pdf*, $f_X(x)$, of a continuous random variable X is the function that satisfies

(1.6.3) $$F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt \quad \text{for all } x.$$

We say $X \sim F_X(x)$ or $X \sim f_X(x)$

is distributed as

Similarly
$$P(a \le X \le b) = \int_a^b f_X(t)\,dt$$

Note t is just a placeholder
$$\int_a^b f_X(t)\,dt = \int_a^b f_X(z)\,dz = \int_a^b f_X(y)\,dy$$
(see appendix)

and
$$P(a < X < b) = \int_a^b f_X(t)\,dt$$

and $a=b=x$
$$P(x \le X \le x) = P(X=x) = \int_x^x f_X(t)\,dt = 0$$

Since we can get the cdf from the pdf or pmf

$F_X = P_X(X \le x)$ for all x

discrete sum $F_X = \sum_k^x f_X(k)$ sum over all values in $S \le x$.

continuous integrate $F_X(x) = \int_{-\infty}^{x} f_X(t)\,dt$

"add up" "point probs" to get $F_X(x) = $ cdf

Or get pdf from $F_X(x)$

$$\frac{d}{dx} F_X(x) = f_X(x)$$

pdf (or pmf) and cdf give same info.
use whichever simpler.

Logistic dist'n continued:

Recall $F_X(x) = \frac{1}{1+e^{-x}}$ and $f_X(x) = \frac{d}{dx} F_X(x) = \frac{e^{-x}}{(1+e^{-x})^2}$ (details appendix)

We can also calculate $P(a \le X \le b)$

$$= F_X(b) - F_X(a)$$

$$= \int_{-\infty}^{b} f_X(x)dx - \int_{-\infty}^{a} f_X(x)dx$$

Note x is now placeholder and upper bound of integral is a or b. (see appendix)

$$= \int_a^b f_X(x)dx$$

Logistic curve (pdf) - cdf is area under pdf curve to point x.

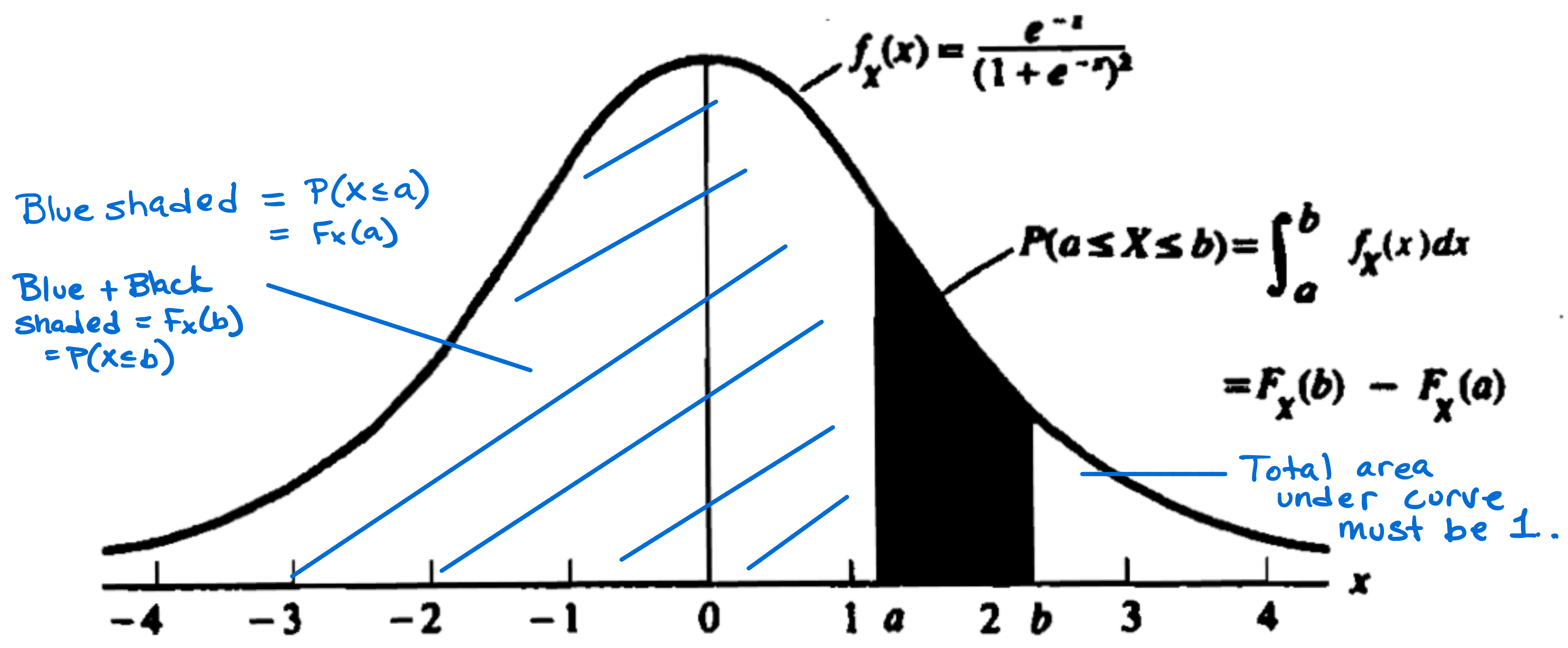


Figure 1.6.1. *Area under logistic curve*

Requirements for pdf (or pmf):

Theorem 1.6.5 *A function $f_X(x)$ is a pdf (or pmf) of a random variable X if and only if*

a. $f_X(x) \ge 0$ *for all x.*

b. $\sum_x f_X(x) = 1$ *(pmf)* *or* $\int_{-\infty}^{\infty} f_X(x)\,dx = 1$ *(pdf).*

know $\lim_{x\to\infty} F_X(x) = 1$ $\lim_{x\to -\infty} F_X(x) = 0$

and continuous: $F_X(x) = \int_{-\infty}^{x} f_X(t)dt$

for this class we will focus on well behaved functions i.e. we will ignore "pathological" cases.

$$1 = \lim_{x\to\infty} F_X(x) = \int_{-\infty}^{\infty} f_X(t)dt$$

Appendix:

Geometric Series Summation Formula: Proof by induction:

Proof by induction: Show for $n=1$ Show $\sum_{k=1}^{n} t^{k-1} = \frac{1-t^n}{1-t}$
Assume for $(n-1)$
Show for n

Page 4

$n=1$ $\sum_{k=1}^{1} t^{k-1} = t^0 = 1$ and $\frac{1-t^1}{1-t} = 1$ //

Assume $\sum_{k=1}^{n-1} t^{k-1} = \frac{1-t^{n-1}}{1-t}$ and show $\sum_{k=1}^{n} t^{k-1} = \frac{1-t^n}{1-t}$

$$\sum_{k=1}^{n} t^{k-1} = \sum_{k=1}^{n-1} t^{k-1} + t^{n-1} = \frac{1-t^{n-1}}{1-t} + t^{n-1} = \frac{1-t^{n-1}+t^{n-1}(1-t)}{(1-t)} = \frac{1-\cancel{t^{n-1}}+\cancel{t^{n-1}}-t^n}{(1-t)}$$

$$= \frac{1-t^n}{(1-t)}$$ // Q.E.D.

Page 5

Show $F_X(x) = 1-(1-p)^x$ is increasing function

$= \frac{d}{dx} 1 - \frac{d}{dx}(1-p)^x$

$= 0 - \frac{d}{dx} \exp(x \log(1-p))$ – by properties of logs

Note: For this class we will use log rather than 'ln' for natural log to avoid confusion w/ n, our sample size (later).

$= -\log(1-p)(1-p) > 0$

$\Rightarrow$ increasing function

– also remember if $0<p<1$ (or $0<(1-p)<1$) $\log(p) < 0$

Page 6, 10

Show $F_X(x) = \frac{1}{1+e^{-x}}$ is increasing in x

$$\frac{d}{dx} F_X(x) = \frac{d}{dx}(1+e^{-x})^{-1} = -1(1+e^{-x})^{-2} * (-e^{-x})$$

$$= \frac{e^{-x}}{(1+e^{-x})^2}$$ num >0 ($e^x>0$ $-\infty<x<\infty$) den >0

$\frac{d}{dx} F_X > 0$
F_X is increasing ft'n.

Note on integration and summation for obtaining cdf from pdfs.

Page 9

pmf: $P(X \le x) = \sum_{i=1}^{x} P(X=i)$
$= \sum_{j=1}^{x} P(X=j)$

i and j are just placeholders here

---

pdf: $P(X \le x) = \int_{-\infty}^{x} f(t)dt$
$= \int_{-\infty}^{x} f(z)dz$

t and z are just placeholders here

we sometimes do this:

$\int_{-\infty}^{x} f(x)dx$, which can get confusing...