# CHAP TWO



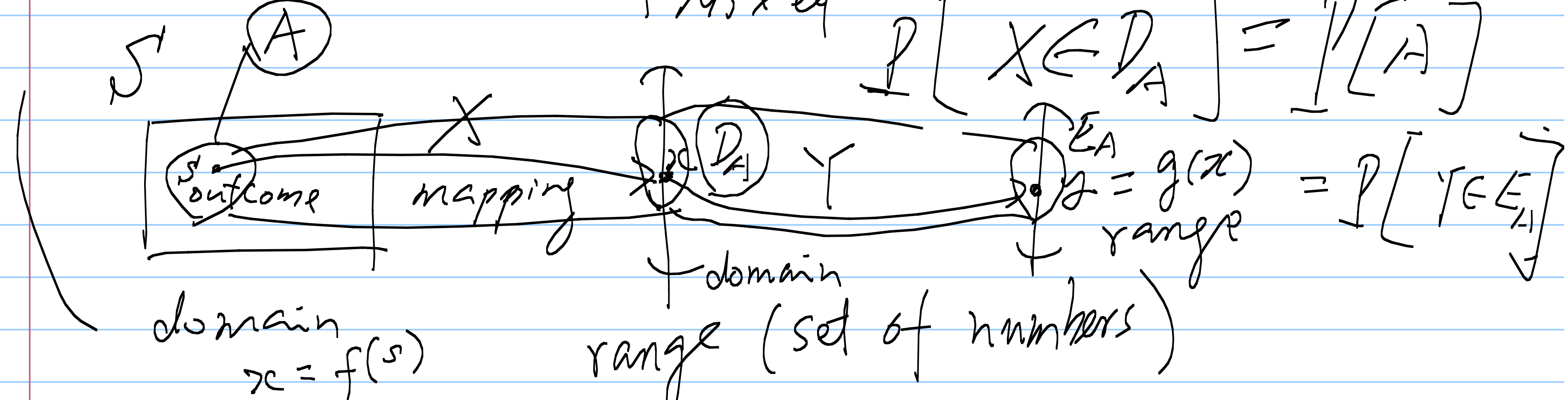
Experiment : Procedure and Observation

Physical Model $\Rightarrow$ Probability Model

Random Variable
- Discrete ✓
- Continuous ✓
- Mixed

$P[A]$

$P[X \in D_A] = P[A] = P[Y \in E_A]$

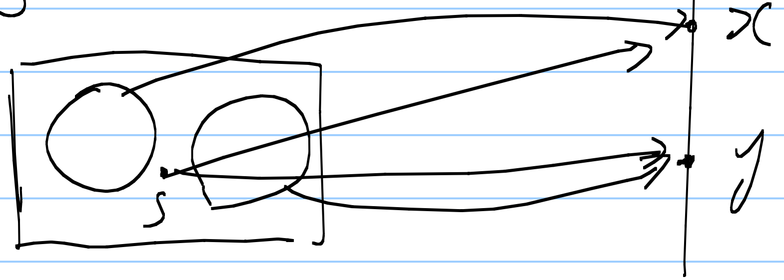


$$\boxed{P[\underbrace{X = x}]} \equiv P[A] \text{ where}$$

$$A = \{s \mid \underbrace{X(s) = x}\}$$

$$P_X(x) \triangleq P[X = x]$$

$S'$ $S_X$

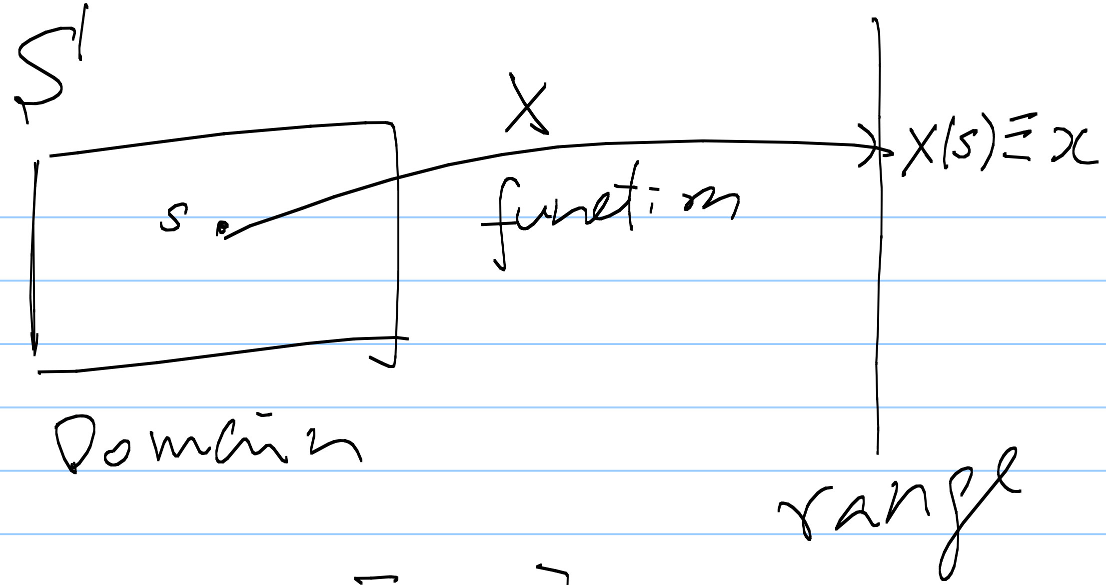




domain range

$$\{X = x\} \cap \{X = y\} = \phi$$

if $x \neq y$.




$X(s)$

$$P_X(x) \triangleq P[X = x] = P[\{s \mid X(s) = x\}]$$

r or a

$$\overbrace{\underbrace{X\ X\ \cdots\ X}_{x-1}\ r}^{x}$$

$k-1$   $r$   $\binom{x-1}{k-1} p^{k-1} (1-p)^{(x-1)-(k-1)} \cdot p$

$x-(k-1)$   $a$

Recall: Consider Repeatable Experiment with identical and independent subexperiments. Each subexperiment gives two possible outcomes, namely success, with probability $p$, and failure, with probability $1-p$.

Bernoulli $(p)$ random variable represents the number of successes in one single trial

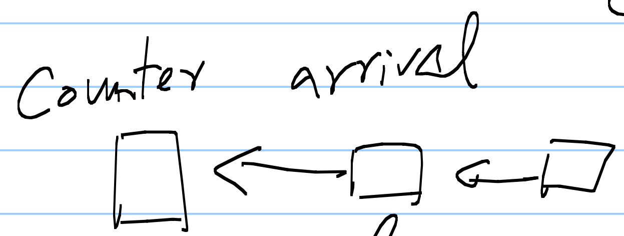
Geometric $(p)$ random variable represents the number of trials until the first success (inclusive)

Binomial $(n, p)$ random variable represents the number of successes in $n$ trials.

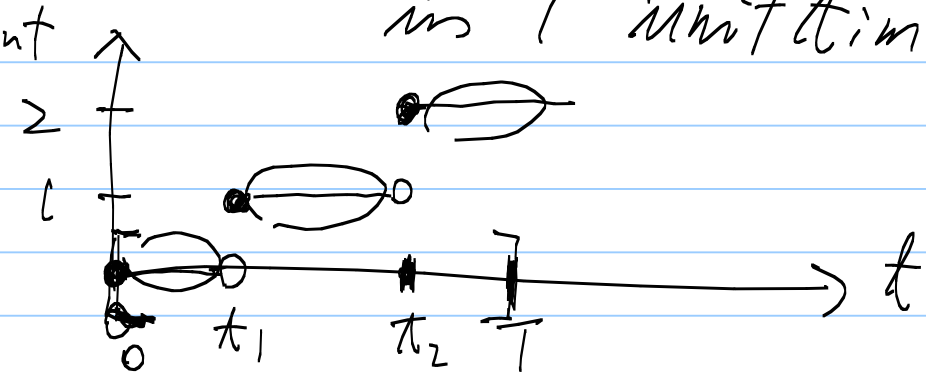
Note: Binomial $(1, p)$ = Bernoulli $(p)$

Pascal $(k, p)$ random Variable represents the number of trials until $k$ successes (inclusive). Note: Pascal $(1, p)$ is equivalent to Geometric $(p)$.

Poisson $(\alpha)$ Random Variable = number of arrivals in $T$ unit times



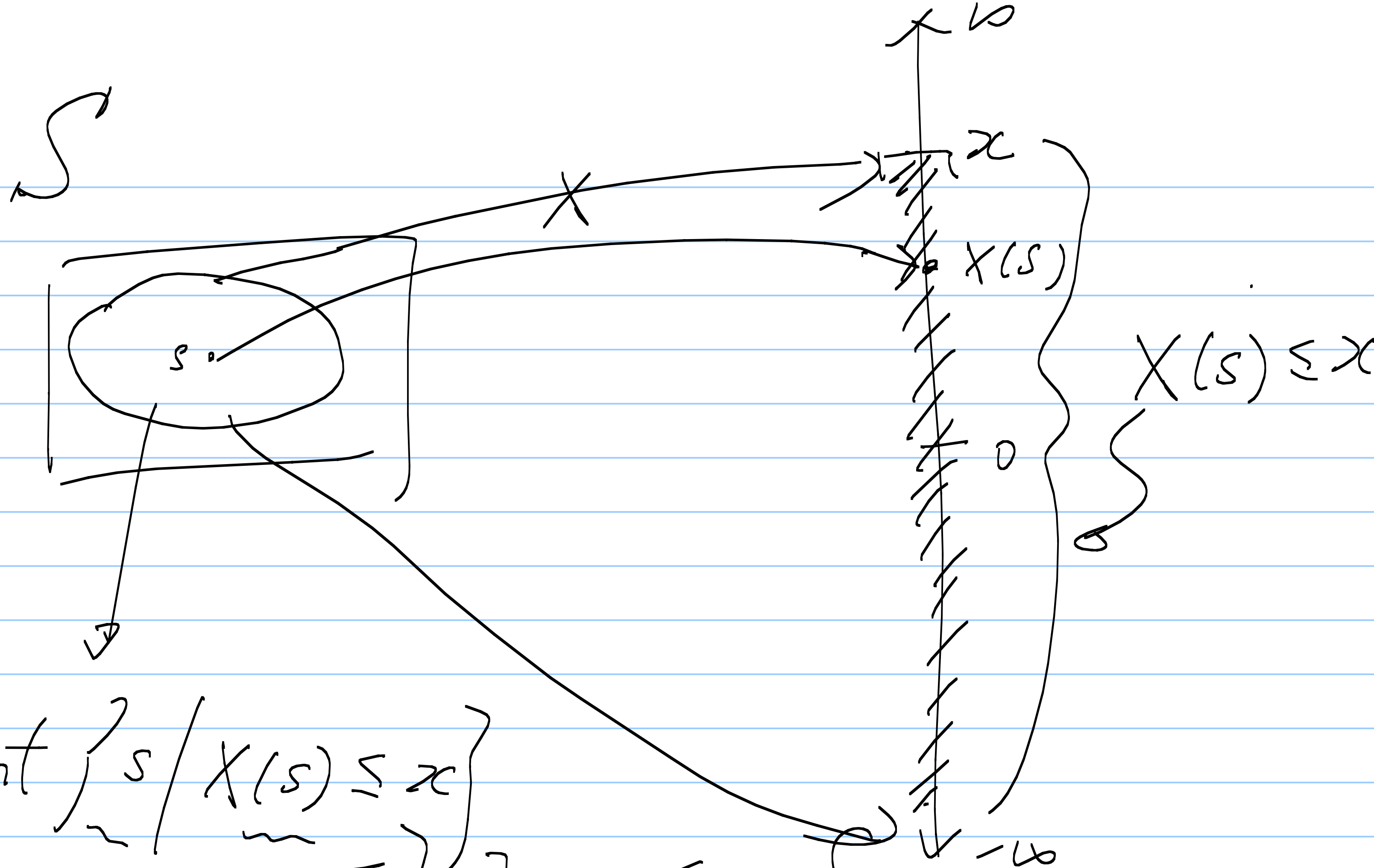

Inter-Arrivals occur independently



$\alpha$ denotes the average number of arrivals in $T$ unit times

$\lambda$ denotes the average number of arrivals in unit time

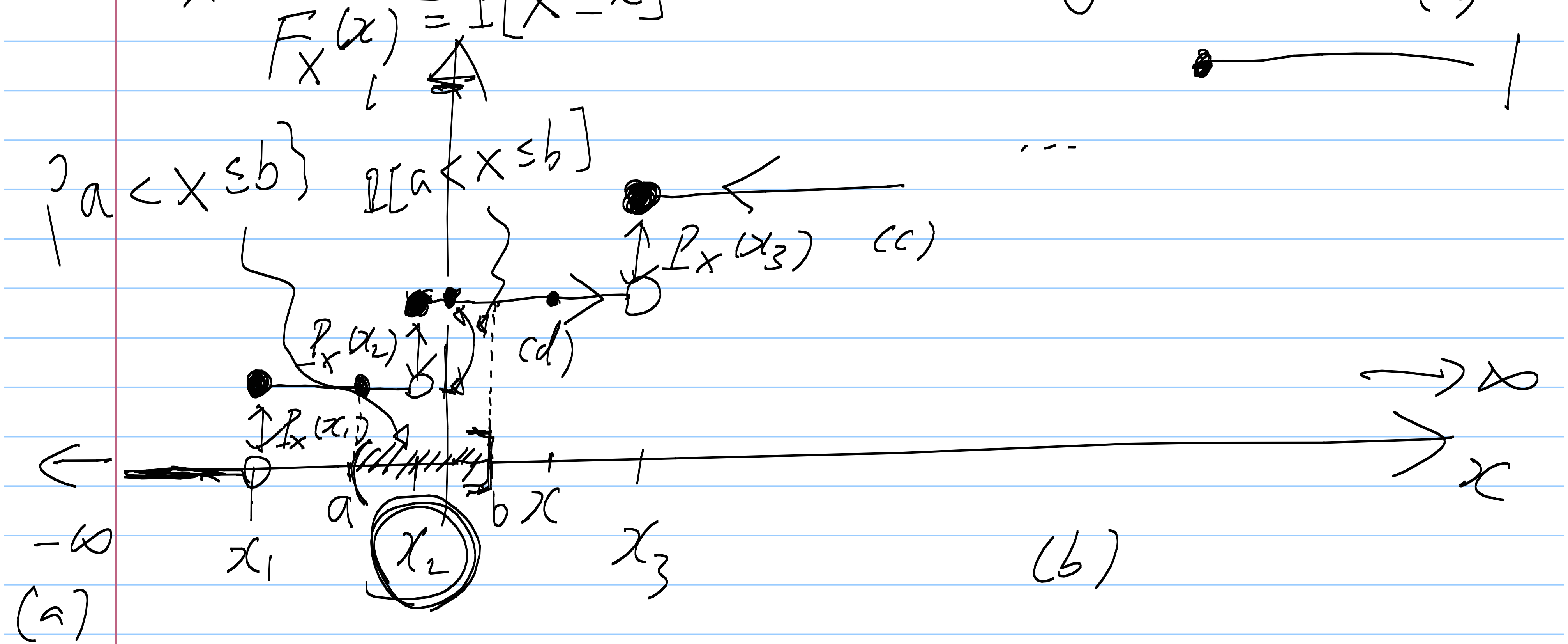
$$\alpha = \lambda T \qquad P_k(k) = \begin{cases} \alpha^k e^{-\alpha}/k! \,, & k = 0, 1, \ldots \\ 0 \,, & \text{otherwise} \end{cases}$$




event $\{ s \mid X(s) \le x \}$

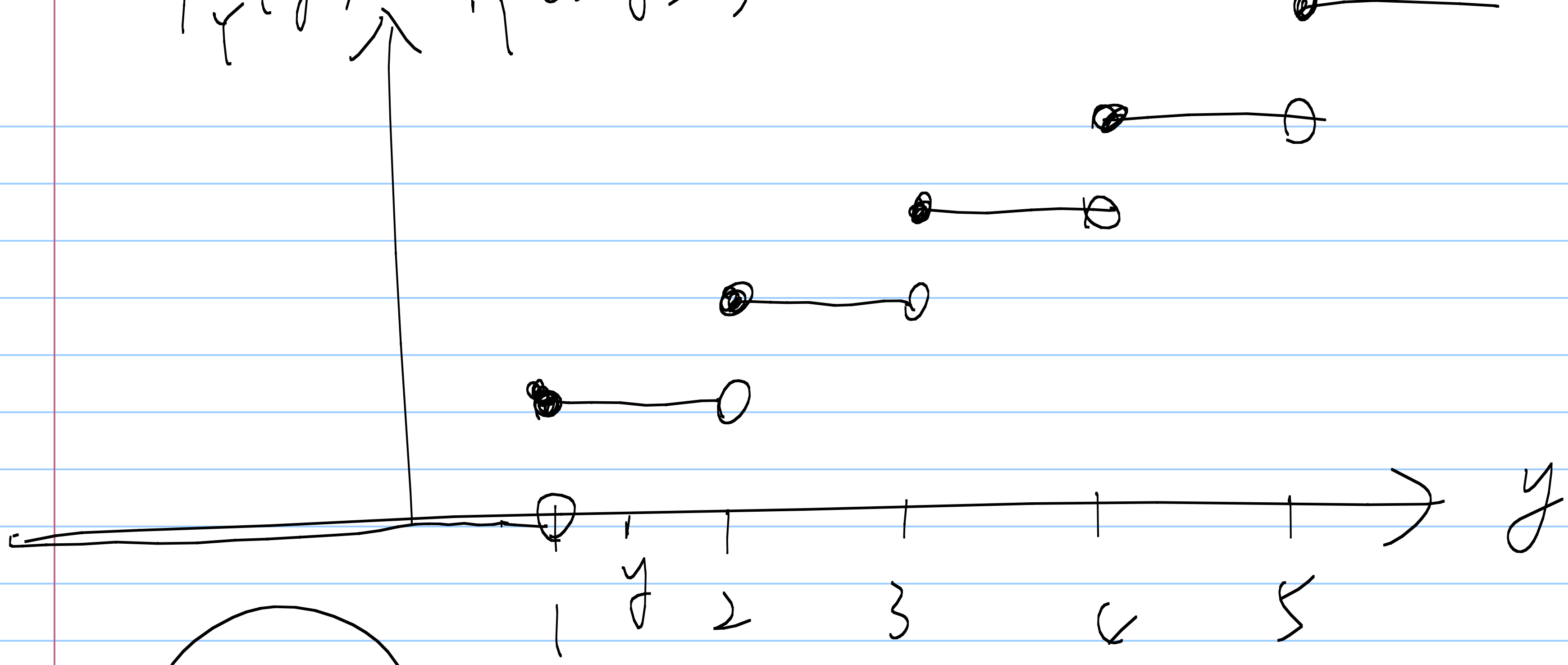
$F_X(x) \triangleq P[\{\ \}] = P[X \le x] \neq P[X < x]$

$F_X(x)$ must be nondecreasing.




(e) $F_X(x)$ is continuous from the right, but not necessarily from the left.

$F_Y(y) = F_Y(\lfloor y \rfloor)$




$\lfloor y \rfloor$ ≡ the largest integer that is not larger than $y$.

≡ integer part of $y$

$\lceil y \rceil \equiv$ the smallest integer that is not smaller than $y$

- The <u>average</u> value of a set of (numerical) experiment outcomes is a statistic of the outcomes.

Three averages of interest:

mean (or expected value, or expectation)

median

mode (unimodal, multimodal)

Consider an experiment that produces a rv $X$. Perform $n$ independent trials of this experiment. Let $x_i$ denote the $i$th sample value.

$$m_n \equiv \text{sample average} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

$$= \frac{1}{n}\sum_{x \in S_X} x N_x$$

( $N_x \equiv$ the number of sample value $x$ in $n$ trials )

$$m_n = \sum_{x \in S_X} x \frac{N_x}{n}$$

We can interpret $P[X = x] = \lim_{n \to \infty} \frac{N_x}{n}$

( relative frequency interpretation )

$$\lim_{n\to\infty} m_n = \lim_{n\to\infty} \sum_{x \in S_X} x \frac{N_x}{n}$$

$$= \sum_{x \in S_X} x \underbrace{\lim_{n\to\infty} \frac{N_x}{n}}_{P[X=x]} = \sum_{x \in S_X} x P[X=x] = E[X]$$

$\checkmark$ $$\sum_{x=0}^{\infty} q^x = \frac{1}{1-q} \quad \text{for } 0<q<1$$

$$\sum_{x=1}^{\infty} x q^x = \sum_{x=0}^{\infty} x q^x = \sum_{x=0}^{\infty} q \frac{d}{dq} q^x$$

$$= q \frac{d}{dq} \underbrace{\sum_{x=0}^{\infty} q^x}_{\frac{1}{1-q}} = q \underbrace{\frac{d}{dq} (1-q)^{-1}}_{+(1-q)^{-2}}$$

$$= \frac{q}{(1-q)^2}$$

$$\Rightarrow \frac{p}{\cancel{q}} \frac{\cancel{q}}{(1-q)^2} = \frac{p}{(1-q)^2} = \frac{1}{p}.$$

$$\lim_{n\to\infty} P_K(k) = \underbrace{\lim_{n\to\infty} \frac{n(n-1)\cdots(n-k+1)}{n^k}}_{=1} \cdot \frac{\alpha^k}{k!} \cdot \underbrace{\lim_{n\to\infty}\left(1-\frac{\alpha}{n}\right)^{n-k}}_{=\frac{\left(1-\frac{\alpha}{n}\right)^n}{\left(1-\frac{\alpha}{n}\right)^k}}$$

$$= \frac{\alpha^k}{k!} e^{-\alpha}$$

$$= \frac{\lim_{n\to\infty}\left(1-\frac{\alpha}{n}\right)^n}{\lim_{n\to\infty}\left(1-\frac{\alpha}{n}\right)^k} \qquad \lim_{n\to\infty}\left(1-\frac{\alpha}{n}\right)^k = 1$$

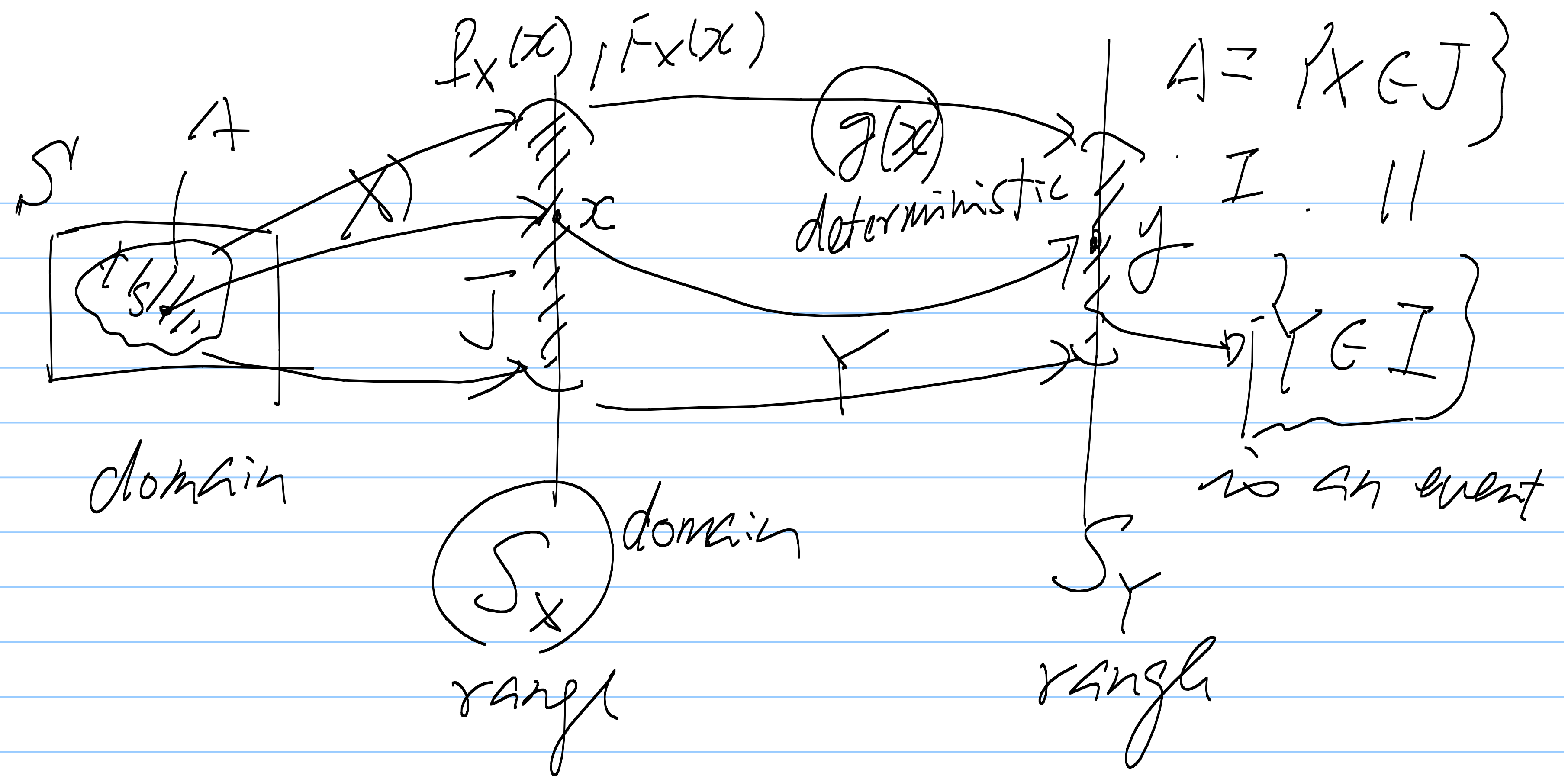
$$= \lim_{n\to\infty}\left(1-\frac{\alpha}{n}\right)^n = \lim_{n\to\infty} e^{n\ln\left(1-\frac{\alpha}{n}\right)}$$

$$\frac{0}{0}$$

$$= e^{\lim_{n\to\infty} n \ln\left(1-\frac{\alpha}{n}\right)} = e^{\lim_{n\to\infty} \frac{\ln(1-\alpha/n)}{1/n}}$$

L'Hopital law

$$= e^{\lim_{n\to\infty} \frac{\frac{\alpha/n^2}{1-\alpha/n}}{-n^{-2}}} = e^{\lim_{n\to\infty} \frac{-\alpha}{1-\alpha/n}} = e^{-\alpha}$$

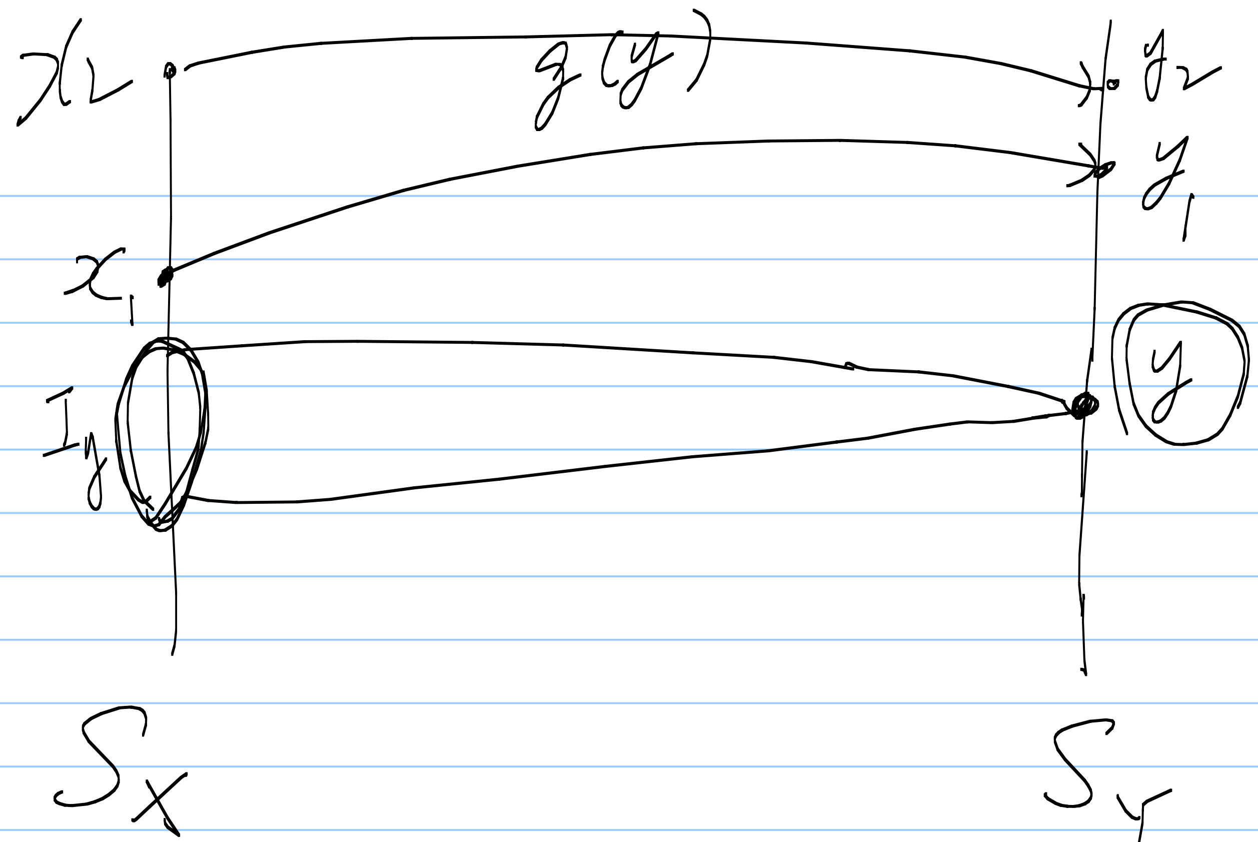



$X$ is completely statistically characterized if $f_X(x)$ or $F_X(x)$ is known.

$A = \{s \mid X(s) \in J\} = \{s \mid Y(s) \in I\}$

$Y$ is called the derived random variable.

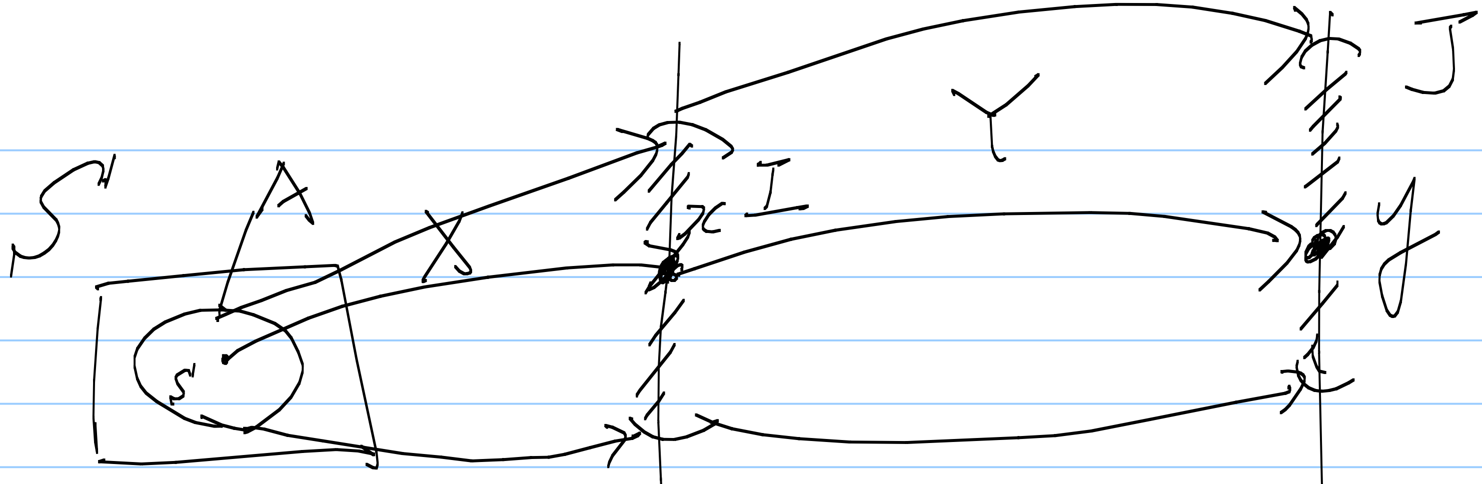
Denote $\boxed{Y = g(X)}$ ✓




① When $g(x_1) \neq g(x_2)$ for $x_1 \neq x_2$ (i.e., $g(x)$ is one-to-one), $P_Y(y) = P_X(x)$ if $y = g(x)$.

② When $y = g(x)$ for all $x \in I_y \triangleq \{x \mid g(x) = y\}$,

$$P_Y(y) = \sum_{x \in I_y} P_X(x)$$




domain

$S_X$ range

$S_Y$ range

$$A = \{ s \mid X(s) \in I \}$$

$$= \{ X \in I \}$$

$$= \{ Y(X(s)) \in J \} = \{ Y \in J \}$$

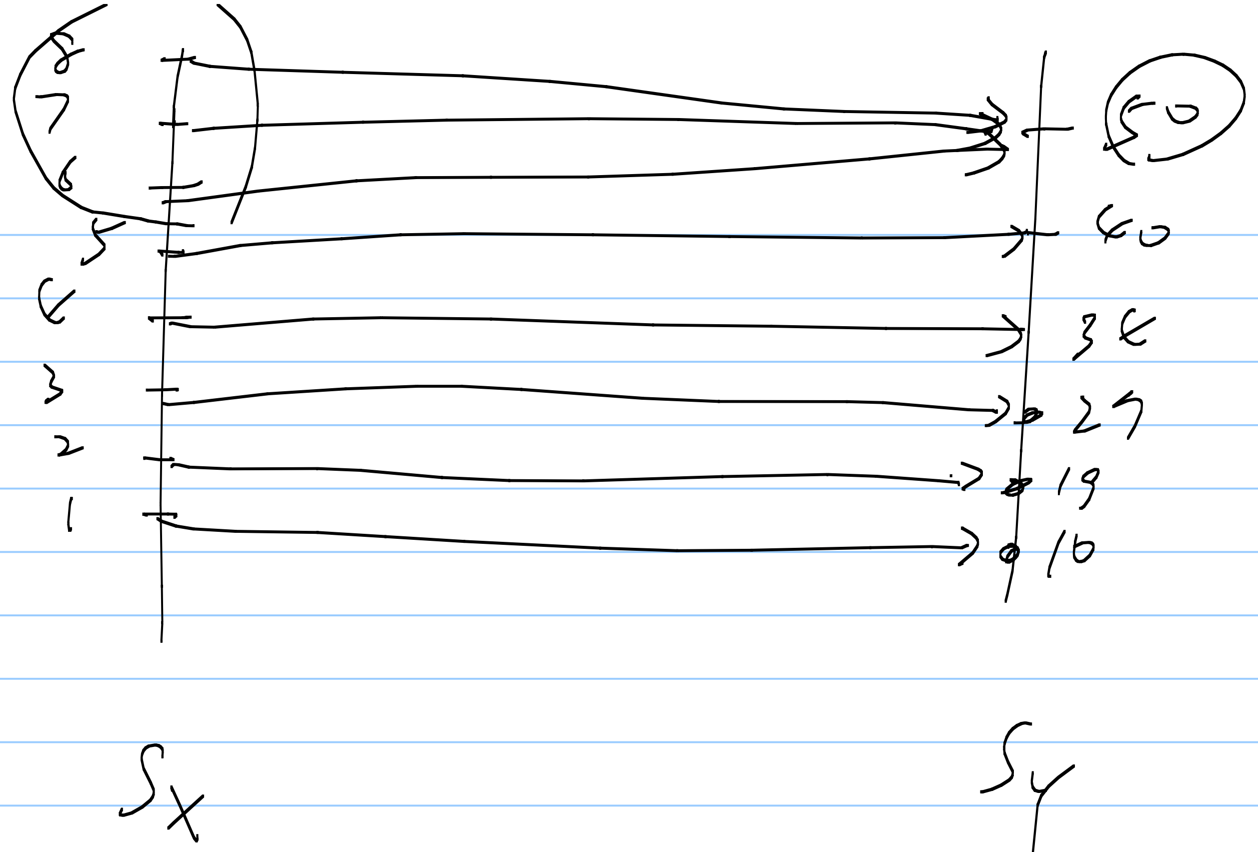
domain

$$Y = g(X)$$

When $J = \{y\}$, $A = \{x \mid g(x) = y\} = \{s \mid g(x(s)) = y\}$ and

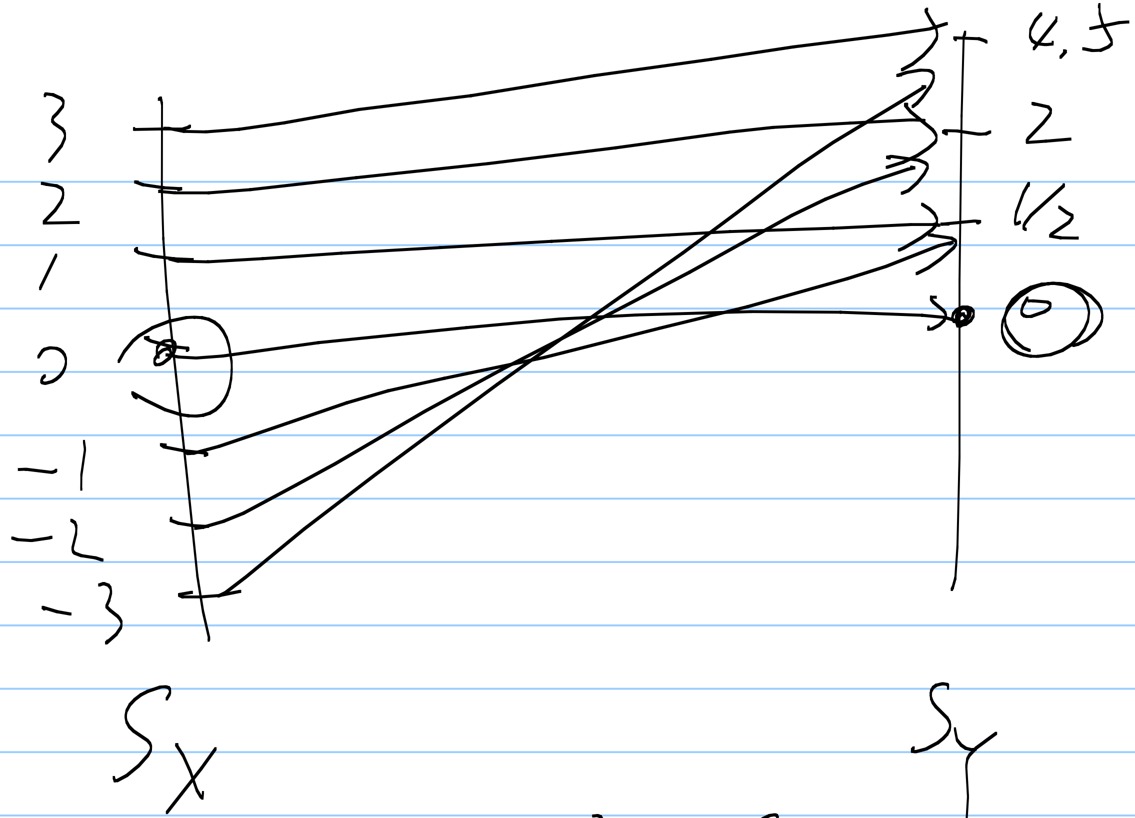
$$P[Y \in J] = P[Y = y] = P[A] = \sum_{x : g(x) = y} P_X(x)$$

$$P[Y = y] = P_Y(y)$$




many-to-one mapping

Given $P_X(x)$ and $y = g(x)$, this section deals with the approach of finding $P_Y(y)$.

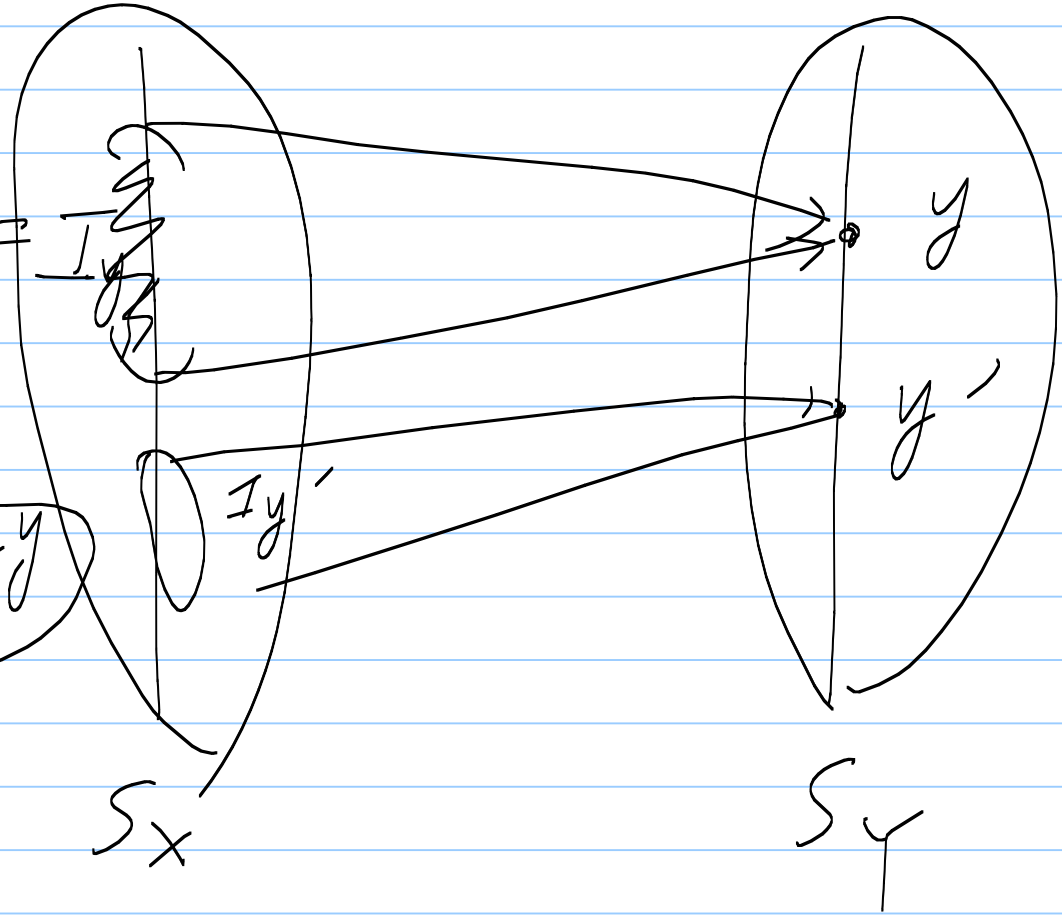



$\{Y=0\} = \{X=0\}$, $\{Y=\frac{1}{2}\} = \{X \in \{-1, 1\}\}$

$\{Y=2\} = \{X \in \{-2, 2\}\}$, $\{Y=4.5\} = \{X \in \{-3, 3\}\}$

# Law of Unconscious Statistician

$\{x \mid g(x)=y\} = I_y$




$$\sum_{y \in S_Y} \sum_{x:\, g(x)=y} = \sum_{x \in S_X}$$

- An operator $\overrightarrow{\Pi}[x]$ is called linear (affine) iff for any $a_n$'s and any $N$, and for any arguments $x_n$'s

$$\overrightarrow{\Pi}\left[\sum_{n=1}^{N} a_n x_n\right] = \sum_{n=1}^{N} a_n \overrightarrow{\Pi}[x_n]$$

(Principle of Superposition).

- For any two linear operators $\overrightarrow{\Pi}[x]$ and

$S[Y]$, $\quad \mathbb{I}[S[Y]] = S[\mathbb{I}[Y]]$

- Expectation is linear.

- pf for Theorem 2.12:

Let $Y = g(X) = aX + b$. Then, from theorem 2.10,

$$E[Y] = \sum_{x \in S_X} g(x) P_X(x) = \underbrace{\sum_{x \in S_X} (ax + b) P_X(x)}$$

$$= a \underbrace{\sum_{x \in S_X} x P_X(x)}_{\mu_X} + b \underbrace{\sum_{x \in S_X} P_X(x)}_{= 1} \quad \text{Q.E.D.}$$

If $W = g(R)$ where $g(r)$ is nonlinear, then

$$E[W] \neq g(E[R])$$

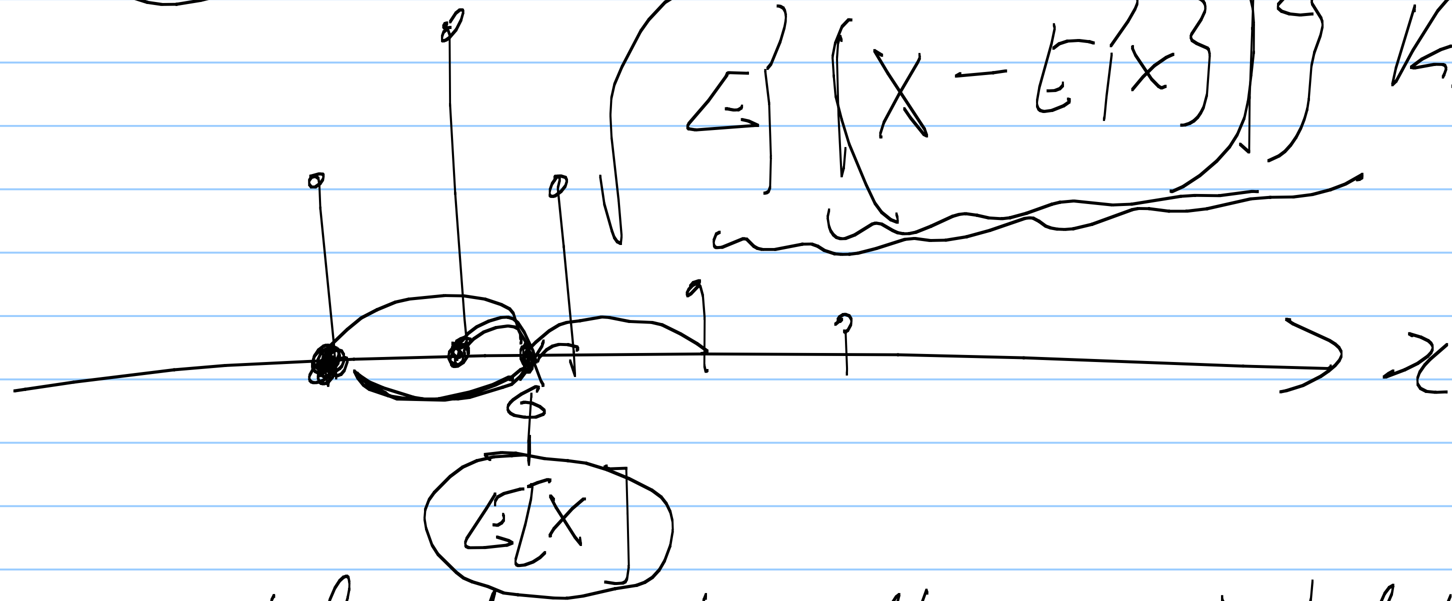
If $W = g(R)$ where $g(r)$ is linear, then

$$E[W] = g(E[R]).$$

$P_X(x)$

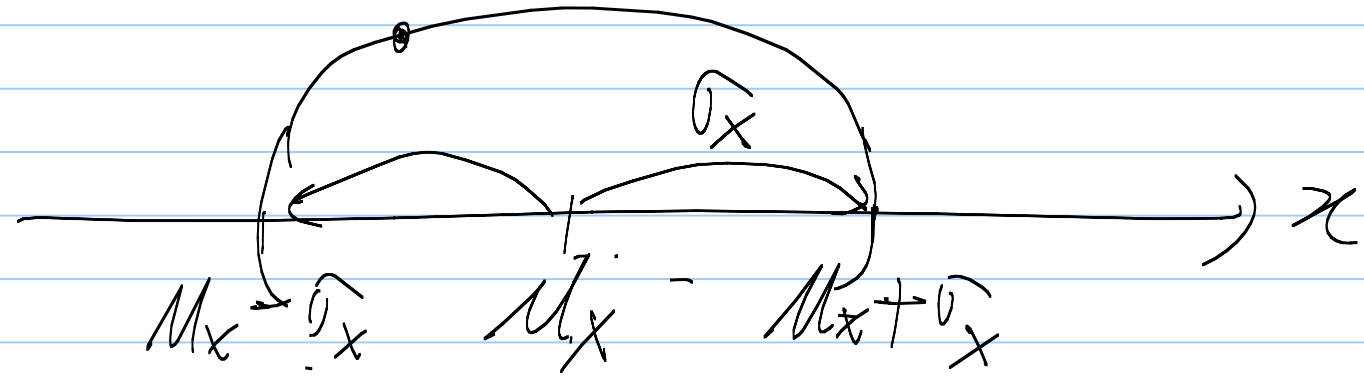
$$E\{X - E[X]\} = 0$$

$$\sqrt{E\{[X - E[X]]^2\}} \quad \text{Variance}$$




$P_X(x)$ completely describe the probability model of $X$, and so does $F_X(x)$.

For some $X$,




$$P\left[ X \in (\mu_X - \sigma_X, \mu_X + \sigma_X) \right]$$

$Var[X]$ or $\sigma_X$ describes the dispersion of $X$ relative to $\mu_X$.

$$Var[X] \triangleq E[\underbrace{(X-\mu_X)^2}]$$

$$= \sum_{x \in S_X} \underbrace{(x-\mu_x)^2}_{(x^2 - 2\mu_x x + \mu_x^2)} P_X(x)$$

law of unconscious statistician

$E[X^2]$ is commonly called the <u>mean square value</u> of $X$.

All the moments $\{E[X^n] \mid n \in \{1, 2, \cdots\}\}$ completely describe the probability model of $X$.

$$Var[X] \triangleq E[(X - \mu_X)^2] \geq 0$$

$$= E[X^2] - \mu_X^2 \geq 0$$

$\Longrightarrow$ $\boxed{E[X^2] \geq \mu_X^2 \text{ for any random variable } X.}$

Recall: $\mathbb{P}[A|B]$ denotes the conditional probability of event $A$ given that event $B$ occurs, with $\mathbb{P}[B] > 0$. That is, we define

$$\mathbb{P}[A|B] = \frac{\mathbb{P}[A \cap B]}{\mathbb{P}[B]}.$$

If we let $A = \{X = x\}$ for a discrete random variable, then

$$\mathbb{P}[A|B] = \mathbb{P}[X = x | B] \equiv p_{X|B}(x).$$

$$\underbrace{P_{X|B}(x)} \triangleq P[X = x \mid B] = \frac{P[X = x, B]}{P[B]}$$

$P_{X|B}(u)$, $P_{X|B}(\cdot)$

dummy argument

If $x \in B$,

$$P_{X|B}(x) = \frac{P[X=x, B]}{P[B]} = \frac{P[X=x, X \in B]}{P[B]}$$

$$= \frac{P[X=x]}{P[B]} = \frac{P_X(x)}{P[B]}$$

If $x \notin B$,

$$P_{X|B}(x) = \frac{0}{P[B]} = 0$$

Q.E.D.

Note that conditional expectation $E[X|B]$ satisfies the law of unconscious statistician, i.e.,

$$E[g(X)|B] = \sum_{x \in B} g(x) \, p_{X|B}(x),$$

and is linear, i.e., for any constants $a, b$,

$$E[aX + b|B] = a\,E[X|B] + b.$$