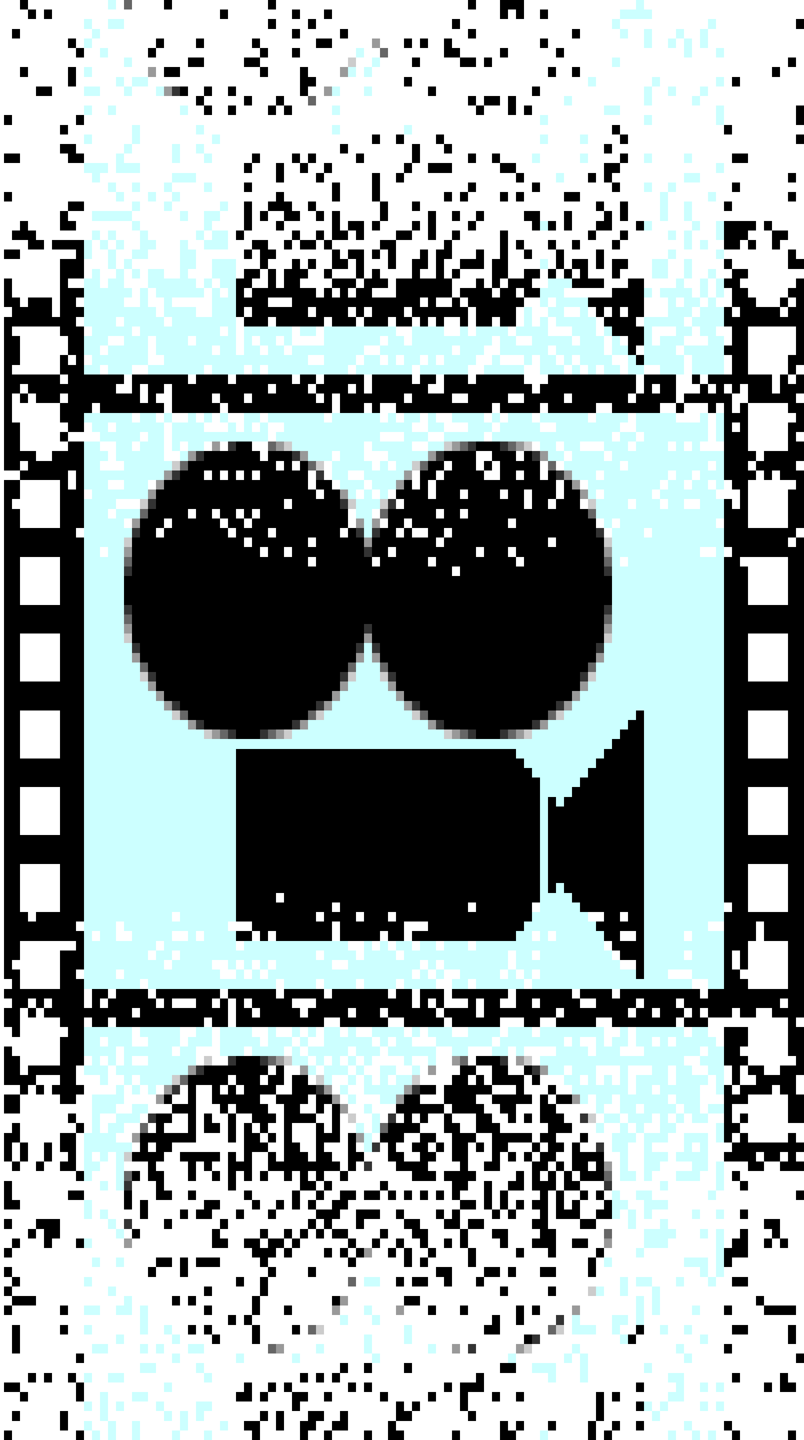
Chapter 1

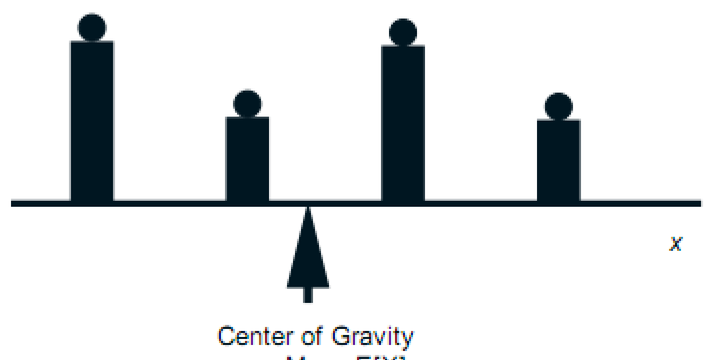
Part 2

Expected Values, Variance and The Properties



Review : MEAN

- The pdf of a RV X provide us with several numbers → the probabilities of all the possible values of X
- Desirable to summarize this information in a single representative number
- Accomplished by the expectation of X → which is a weighted (in proportion to probabilities) average of the possible values of X → the center of gravity of the pdf



Mean & Variance of a discrete R.V

- Mean → describe the "center" of the distribution of X in manner similar to the balance point of a loading
- Variance→ measure of dispersion or scatter in the possible values for X

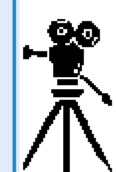
The **mean or expected value** of the discrete random variable X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \sum_x x f(x) \qquad (3\text{-}3)$$

The **variance** of X, denoted as σ^2 or $V(X)$, is

$$\sigma^2 = V(X) = E(X-\mu)^2 = \sum_x (x-\mu)^2 f(x) = \sum_x x^2 f(x) - \mu^2$$

The **standard deviation** of X is $\sigma = \sqrt{\sigma^2}$.



Illustration

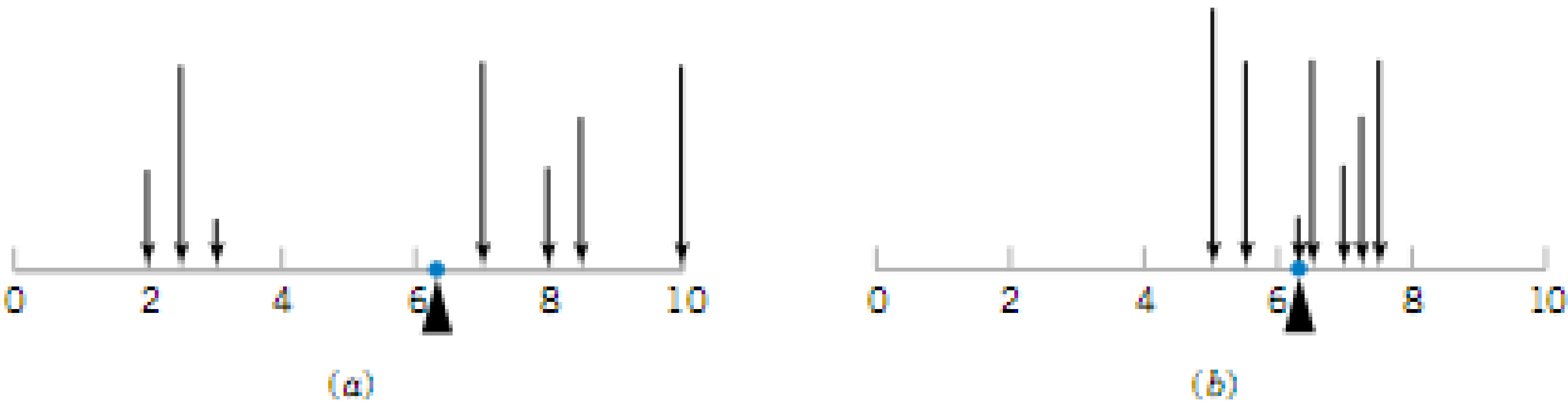


Figure 3-5 A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but Part (a) illustrates a larger variance.

PROVE it

If X is a discrete random variable with probability mass function $f(x)$,

$$E[h(X)] = \sum_x x h(x) f(x) \qquad (3\text{-}4)$$

Example

Consider 2 independent coin tosses, each with a ¾ probability of a head. And let X be the number of heads obtained. the pdf :

$$f(x) = \begin{cases} \left(\frac{1}{4}\right)^2, x = 0 \\ 2.\left(\frac{1}{4}\right).\left(\frac{3}{4}\right), x = 1 \\ \left(\frac{3}{4}\right)^2, x = 2 \end{cases}$$

$$E[X] = (0)\left(\frac{1}{4}\right)^2 + (1)2\left(\frac{1}{4}\right)\left(\frac{3}{4}\right) + (2)\left(\frac{3}{4}\right)^2$$

$$= \frac{24}{16} = \frac{3}{2}$$

Review : Variance

- The variance → a measure of dispersion of X around its mean
- Other measure is standard deviation σ

$$\mathrm{var}(X) = E\left[(X - E[X])^2\right]$$

$$\sigma_x = \sqrt{\mathrm{var}(X)}$$

Exercise 1

1. Let Y=|X| and and suppose :

$$f(x) = \begin{cases} \frac{1}{9}, & \text{if } x \text{ is an integer in the range [-4,4]} \\ & 0, \text{otherwise} \end{cases}$$

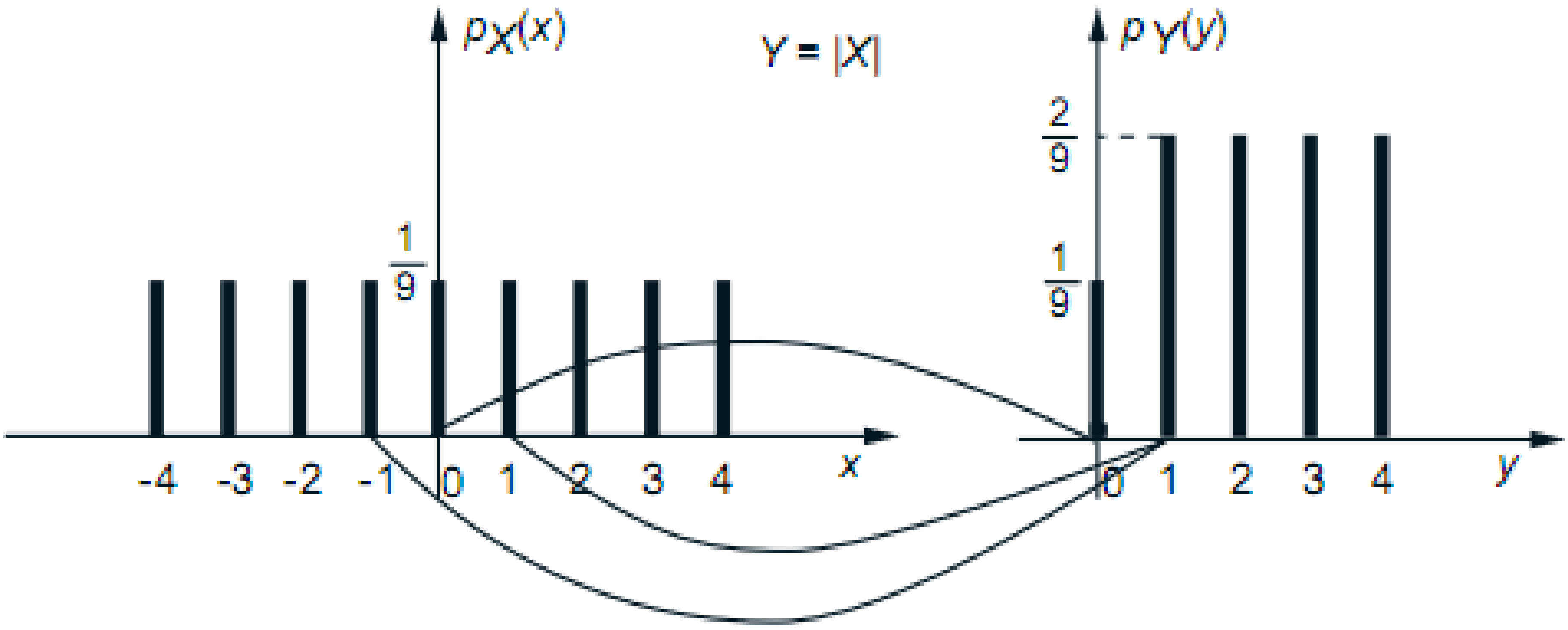
Then find y! and f(y)!

2. What about if

$$Z = X^2$$

Then find z! and f(z)!

ilustration



Exercise 2

Look at Exercise 1

1. Find E[X], E[Z] !
2. Var (X)

Exercise

1. For a discrete random variable Y the probability distribution is

$$p(y) = \frac{5-y}{10} \quad \text{for} \quad y = 1,\ 2,\ 3,\ 4.$$

Calculate: (a) $E(Y)$ (b) $V(Y)$.

2. For a fair 10-sided spinner, if S is 'the score on the spinner', find:

(a) the probability distribution of S; (b) $E(S)$;

(c) the standard deviation of S.

3. A random variable has probability distribution

x	0	1	2	3
$P(X = x)$	0.4	0.3	0.2	0.1

Find:

(a) the mean and variance of X;

(b) the mean and variance of the random variable

$$Y = X^2 - 2X.$$

4. A fair six-sided die has

'1' on one face
'2' on two of its faces
'3' on the remaining three faces.

The die is thrown twice, and X is the random variable 'total score thrown'. Find

(a) the probability distribution;

(b) the probability that the total score is more than 4;

(c) $E(X)$ and $V(X)$.

Th.

Mean and Variance of a Linear Function of a Random Variable

Let X be a random variable and let

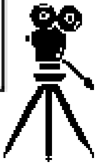
$$Y = aX + b,$$

where a and b are given scalars. Then,

$$\mathbf{E}[Y] = a\mathbf{E}[X] + b, \qquad \mathrm{var}(Y) = a^2 \mathrm{var}(X).$$

Variance in Terms of Moments Expression

$$\mathrm{var}(X) = \mathbf{E}[X^2] - \big(\mathbf{E}[X]\big)^2.$$



1. Measures of Central Tendency (location Parameter)

a. Arithmetic Mean (Mean, Average, Expectation, Expected Value). The mean is the most often used measure of central tendency. It is defined as:

$$\mu = \int_{-\infty}^{\infty} x f(x) dx \quad . \tag{51}$$

b. Median (Introduced in 1883 by Francis Galton). The median m is the value that divides the total distribution into two equal halves, i.e.,

$$F(m) = \int_{-\infty}^{m} f(x) dx = 1/2 \quad . \tag{52}$$

c. Mode (Thucydides, 400 B.C., Athenian Historian). This is also called the most probable value, from the French "mode," meaning "fashion." It is given by the maximum of the distribution, i.e., the value of x for which:

$$\frac{df(x)}{dx} = 0 \quad . \tag{53}$$

2. Measures of Dispersion

(a) Variance (Introduced by R.A. Fisher in 1918):

$$\sigma^2 = \int (x-\mu)^2 f(x)dx = \int x^2 f(x)dx - \mu^2 = E(x^2) - \{E(x)\}^2 \quad . \tag{54}$$

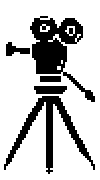
(b) Standard deviation (Introduced by K. Pearson in 1890): The standard deviation is simply the positive square root of the variance, denoted by σ.

(c) Mean deviation:

$$\sigma_d = \int |x-\mu| f(x)dx \quad . \tag{55}$$

(d) Coefficient of variation (c.o.v.): The c.o.v. gives the standard deviation relative to the mean μ as:

$$\text{c.o.v.} = \sigma/\mu \quad . \tag{56}$$



Variance : Properties

if X is a r.V , a and b are constants then :

$$V\left(aX+b\right)=a^2V(X)$$

3. Quantile (we not discussing this!)

4. Higher Moment

(a) Skewness:

$$\alpha_3 = \mu_3/\sigma^3, \text{ where } \mu^3 = \int (x-\mu)^3 f(x)dx \quad . \tag{58}$$

A distribution has a positive skewness (is positively skewed) if its long tail is on the right and a negative skewness (is negatively skewed) if its long tail is on the left.

(b) Kurtosis:

$$\alpha_4 = \mu_4/\sigma^4, \text{ where } \mu_4 = \int (x-\mu)^4 f(x)dx \quad . \tag{59}$$

Kurtosis measures the degree of "peakedness" of a distribution, usually in comparison with a normal distribution, which has the kurtosis value of 3.

(c) Moments of kth order:

Moments about the origin (raw moments):

$$\mu_k' = E(x^k) = \int x^k f(x)dx \quad . \tag{60}$$

Moments about the mean (central moments):

$$\mu_k = E(x-\mu)^k = \int (x-\mu)^k f(x)dx \quad . \tag{61}$$

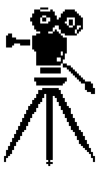
Definition

The k-th moment about the origin
a rV X is :

$$\mu'_k = E[X^k]$$

the k-th moment about the mean

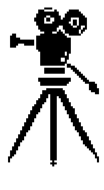
$$\mu_k = E\left[X - E(X)\right]^k = E\left(X - \mu\right)^k$$



Th. Bound on Probability

If X is a RV and u(x) is a nonnegative real valued function, then for any positive constant c>0,

$$P\left[u(X) \geq c\right] \leq \frac{E\left[u(X)\right]}{c}$$



Th. Chebychev Ineq

If X is a RV with mean μ and variance σ^2 then for any k>0,

$$P\left[\left|X-\mu\right| \geq k\sigma\right] \leq \frac{1}{k^2}$$

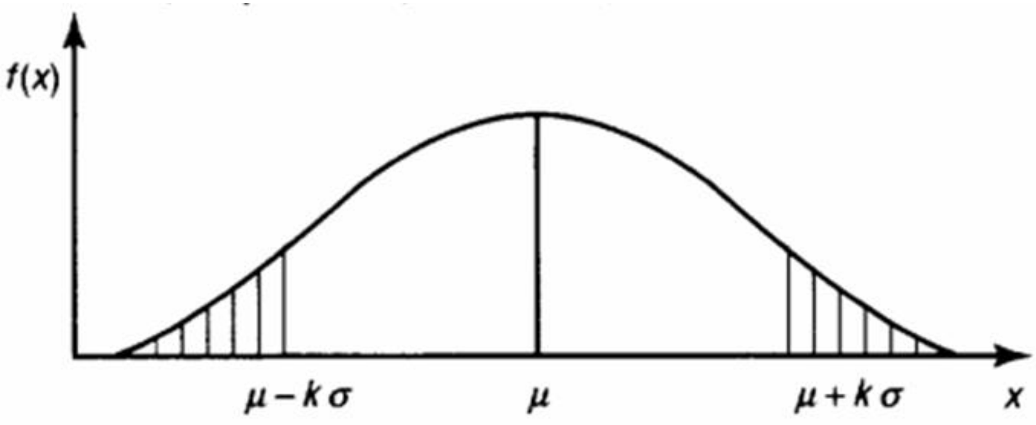


FIGURE 11.—Chebyshev's theorem.

Definition

If X a rV then $M_X(t) = E\left[e^{tX}\right]$

is called the moment generating function (mgf) of X if this expected value exists for all values of t in some interval of the form $-h<t<h$, $h>0$

Theorem

If $Y = aX + b$ then $M_Y(t) = e^{bt} M_X(at)$