Ch4: Distribution Functions and Discrete Random Variables

4.1 Random variables

- In rolling two fair dice, *X* is the sum, then *X* can only assume the values 2, 3, 4, . . . , 12 with the following probabilities:
  - P(X = 2) = P({(1, 1)})= 1/36,
  - P(X = 3) = P({ (1, 2), (2, 1) })= 2/36,
  - P(X = 4) = P({ })=

- and, similarly,

Sum, i	5	6	7	8	9	10	11	12
P(X=i)	4/36	5/36		5/36	4/36	3/36	2/36	

Definition

- *Let S be the sample space of an experiment.*
- *A **random variable** X is a **function** that assigns a real value to each outcome in S.*
- *For any **set of real numbers A**, the prob. that X will assume a value that is contained in A is equal to the prob. that the outcome of the experiment is contained in $X^{-1}(A)$.*
- *That is, $P\{X \in A\} = P(X^{-1}(A))$, where $X^{-1}(A)$ is the event consisting of all points $s \in S$ such that $X(s) \in A$.*

Definition (cont.)

In the previous fair dice rolling example,

- $X(\{(1,1)\})=2$, $X(\{(1,2)\})=3$, $X(\{(2,1)\})=3$, *etc.*
- *If* $A=\{3,4\}$
- $X^{-1}(A)=\{(1,2),(2,1),(1,3),(2,2),(3,1)\}$
- $P\{X \in A\}=P(X^{-1}(A))=2/36+3/36=5/36$

Example 4.3

- In the United States, the number of twin births is approximately 1 in 90.
- Let X be the number of births in a certain hospital until the first twins are born. X is a random variable.
- Denote twin births by T and single births by N. Then X is a real-valued function defined on the sample space

$$S = \{T, NT, NNT, NNNT, \ldots\} \text{ by } X(\underbrace{NNN \ldots N}_{i-1}T) = i$$

The set of all possible values of X is $\{1, 2, 3, \ldots\}$ and

$$P(X = i) =$$

- Let X and Y be two random variables over the same sample space S; then and are real-valued functions having the same domain.
- Therefore, we can form the functions $X + Y$; $X - Y$; $aX + bY$, where a and b are constants; XY; and X/Y, where $Y \neq 0$.
- Similarly, functions such as X^2, $\sin X$, $\cos X^2$, e^X, and $X^3 - 2X$ are random variables.

Example 4.5

- The diameter of a flat metal disk manufactured by a factory is a random number between 4 and 4.5.
- What is the probability that the area of such a flat disk chosen at random is at least 4.41п?

1. D: the diameter of the metal disk selected at random
2. $P(\frac{\pi D^2}{4} > 4.41\pi) = P(D^2 > 17.64) = P(D > 4.2)$
3. the length of *D* is a random number in the interval (4, 4.5) => $P(D > 4.2) =$
4. $P(\frac{\pi D^2}{4} > 4.41\pi) = \frac{3}{5}$

Example 4.6

- A random number is selected from the interval (0, π/2). What is the probability that its sine is greater than its cosine?

1. X：the selected number
2.

$$P(\sin X > \cos X) = P(\tan X > 1) = P(X > \frac{\pi}{4}) =$$

4.2 Distribution functions

- Usually, when dealing with a random variable X, for constants a and b ($b < a$), computation of one or several of the probabilities
  - $P(X = a)$, $P(X < a)$, $P(X \leq a)$, $P(X > b)$, $P(X \geq b)$, $P(b \leq X \leq a)$, $P(b < X \leq a)$, $P(b \leq X < a)$, and $P(b < X < a)$

  is our ultimate goal.

Definition

- *If X is a random variable, then the function F defined on $(-\infty,+\infty)$ by $F(t) =$ is called the **distribution function** of X.*
- Since F "accumulates" all of the probabilities of the values of X up to and including t, sometimes it is called the **cumulative distribution function ( ) of** X.

Properties

1\. F is nondecreasing; that is, if $t < u$, then $F(t) \quad F(u)$.

1\. $\text{Lim}_{t \to \infty} F(t) = 1$.

2\. $\text{Lim}_{t \to -\infty} F(t) = 0$.

3\. F is right continuous. That is, for every $t \in$ **R** , $F(t+) = F(t)$. This means that if t_n is a decreasing sequence of real numbers converging to t , then

$\lim_{n \to \infty} F(t_n) =$

Distribution functions

Event concerning X	Probability of the event in terms of F	Event concerning X	Probability of the event in terms of F
$X \leq a$	$F(a)$	$a < X \leq b$	
$X > a$		$a < X < b$	$F(b-) - F(a)$
$X < a$	$F(a-)$	$a \leq X \leq b$	$F(b) - F(a-)$
$X \geq a$	$1 - F(a-)$	$a \leq X < b$	$F(b-) - F(a-)$
$X = a$			

Example 4.7

- The distribution function of a random variable X is given by

$$F(x)=\begin{cases} 0 & x<0 \\ x/4 & 0\le x<1 \\ 1/2 & 1\le x<2 \\ \dfrac{1}{12}x+\dfrac{1}{2} & 2\le x<3 \\ 1 & x\ge 3 \end{cases}$$

where the graph of F is shown in Figure 4.1. Compute the following quantities:

(a) $P(X < 2)$**; (b)** $P(X = 2)$**; (c)** $P(1 \le X < 3)$**;**
(d) $P(X > 3/2)$**; (e)** $P(X = 5/2)$**; (f)** $P(2 < X \le 7)$**.**

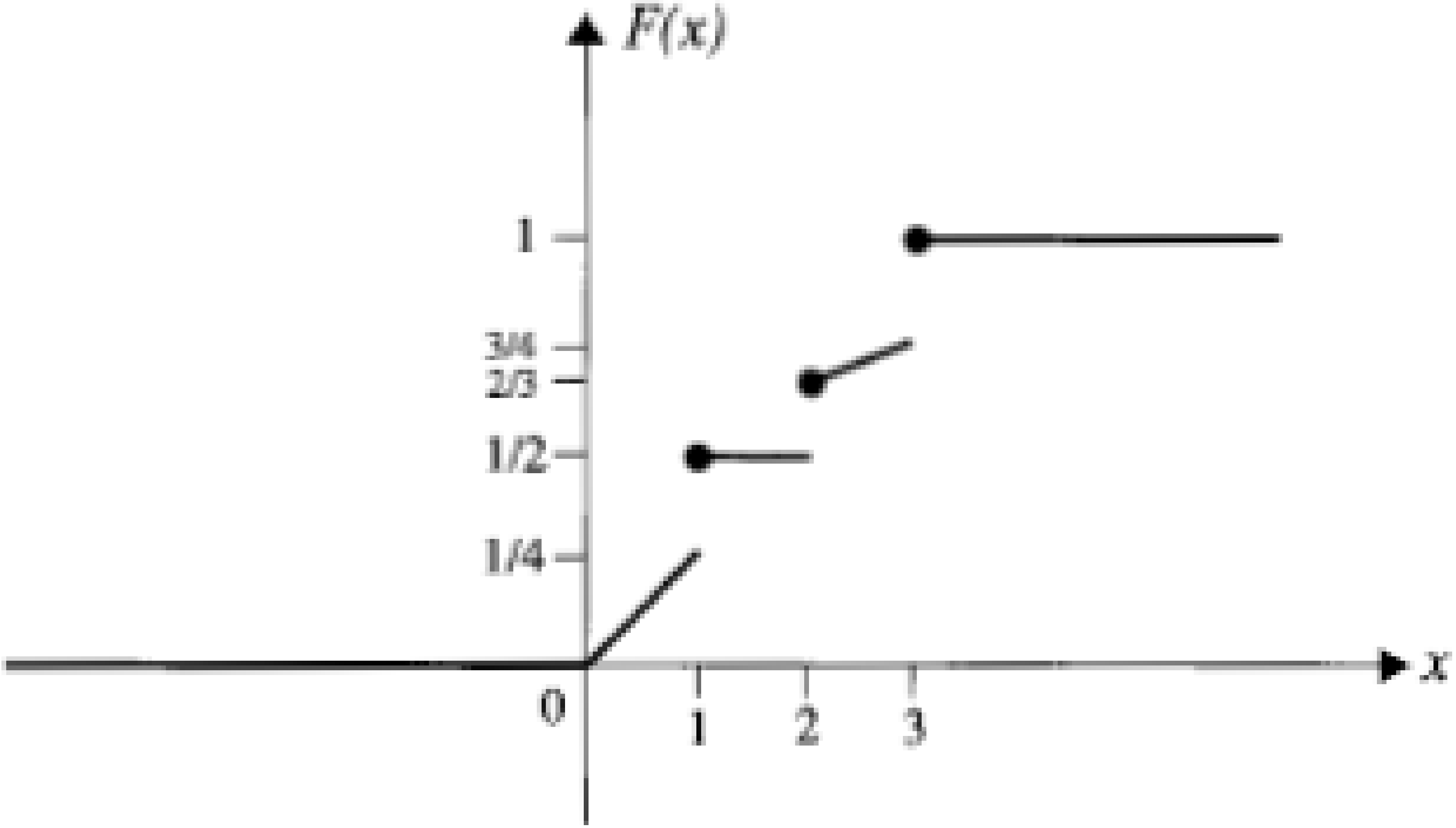

Figure 4.1 Distribution function of Example 4.7.

$$(a) P(X < 2) = F(2-) = 1/2$$

$$(b) P(X = 2) = \qquad\qquad\qquad = (2/12 + 1/2) - 1/2 = 1/6$$

$$(c) P(1 \le X < 3) =$$

$$= (3/12 + 1/2) - 1/4 = 1/2$$

$$(d) P(X > 3/2) = 1 - F(3/2) = 1 - 1/2 = 1/2$$

$$(e) P(X = 5/2) =$$

since F is continuous at 5/2 and has no jumps.

$$(f) P(2 < X \le 7) = F(7) - F(2) = 1 - (2/12 + 1/2) = 1/3$$

Example 4.10

- The sales of a convenience store on a randomly selected day are *X* thousand dollars, where *X* is a random variable with a distribution function of the following form:

$$F(t)=\begin{cases} 0 & t<0 \\ (1/2)t^2 & 0\le t<1 \\ k(4t-t^2) & 1\le t<2 \\ 1 & t\ge 2 \end{cases}$$

Suppose that this convenience store's total sales on any given day are less than $2000.

(a) Find the value of *k*.

(b) Let *A* and *B* be the events that tomorrow the store's total sales are between 500 and 1500 dollars, and over 1000 dollars, respectively. Find *P(A)* and *P(B)*.

(c) Are *A* and *B* independent events?

(a) Since $X < 2$, we have that P(X < 2) = 1,

so F(2−) = 1. This gives k(8−4) = 1, so k =

(b)

$$P(A) = P(\frac{1}{2} \le X \le \frac{3}{2}) = F(\frac{3}{2}) - F(\frac{1}{2}-)$$

$$= F(\frac{3}{2}) - F(\frac{1}{2}) = \frac{15}{16} - \frac{1}{8} = \frac{13}{16}$$

$$P(B) = P(X > 1) =$$

(c)

$$P(AB) = P(1 < X \le \frac{3}{2}) =$$

Since P(AB) ≠ P(A)P(B), *A* and *B* are not independent.

Remark 4.1

- Suppose that F is a right-continuous, nondecreasing function on $(-\infty,\infty)$ that satisfies $\lim_{t\to\infty} F(t) =$ and $\lim_{t\to-\infty} F(t) =$
- It can be shown that there exists a sample space S with a probability function and a random variable X over S such that the distribution function of X is F.
- Therefore, a function is a distribution function if it satisfies the conditions specified in this remark.

4.3 Discrete random variables

- Number of tails in flipping a coin twice
  - finite set
- Number of flips of until the first heads
  - countable set
- Amount of next year's rainfall
  - uncountable set
- Whenever the set of possible values that a random variable X can assume is at most countable, X is called **discrete**.

Definition

- *The* ***probability mass function (****also called* ***probability function*** *or* ***discrete probability function)*** *p of a random variable X whose set of possible values is* $\{x_1, x_2, x_3, \ldots\}$ *is a function from* **R** *to* **R** *that satisfies the following properties.*

(a) $p(x) = 0$ *if* $x \notin \{x_1, x_2, x_3, \ldots\}$.

(b) $p(x_i) = P(\qquad)$

and hence $p(x_i) \geq 0$ $(i = 1, 2, 3, \ldots)$.

(c) $\sum_{i=1}^{\infty} p(x_i) =$

Example 4.12

- Can a function of the form

$$p(x) = \begin{cases} c\left(\frac{2}{3}\right)^x & x = 1,2,3,\ldots \\ 0 & elsewhere \end{cases}$$

be a probability mass function?

1. A probability mass function should have three properties:

 (1) p(x) must be zero at all points except on a finite or countable set. (It is satisfied.)

 (2) p(x) should be nonnegative. This is satisfied if and only if

Example 4.12

(3) $\sum p(x_i) = 1$, This condition is satisfied if and only if . This happens precisely when

$$c = \frac{1}{\sum_{i=1}^{\infty} (2/3)^i} = \frac{1}{\frac{2/3}{1-2/3}} = \frac{1}{2}$$

where the second equality follows from the geometric series theorem. Thus only for c=1/2, a function of the given form is a probability mass function.

Example 4.13

- Let *X* be the number of births in a hospital until the first girl is born. Determine the probability mass function and the distribution function of *X*. Assume that the probability is 1/2 that a baby born is a girl.

1. X is a random variable that can assume any positive integer i. p(i) = P(X=i), and X = i occurs if the first i - 1 birth are all boys and the ith birth is a girl. Thus $p(i) =$ for i = 1, 2, 3,…, and p(x) = 0 if x $\neq$ 1, 2, 3,…

2. To determine $F(t)$. In general for $n-1 \le t < n$

Example 4.13

$$F(t) =$$

$$= \frac{1-(1/2)^n}{1-1/2} - 1 = 1 - \left(\frac{1}{2}\right)^{n-1}$$

by the partial sum formula for geometric series. Thus

$$F(t) = \begin{cases} 0 & t < 1 \\ 1-(1/2)^{n-1} & n-1 \le t < n, n = 2,3,4,\cdots \end{cases}$$

4.4 Expectations of discrete random variables

- **Definition** *The* ***expected value*** *of a discrete random variable X with the set of possible values A and probability mass function p(x) is defined by*

$$E(X) =$$

  We say that E(X) exists if this sum converges absolutely.

- The expected value of a random variable *X* is also called the **mean**, or the **mathematical expectation**, or simply the **expectation** of *X*. It is also occasionally denoted by $E[X]$, EX, μ_X, or μ.

Example 4.17

- In the lottery of a certain state, players pick six different integers between 1 and 49, the order of selection being irrelevant.
- The lottery commission then selects six of these numbers at random as the winning numbers.
- A player wins the grand prize of $1,200,000 if all six numbers that he has selected match the winning numbers. He wins the second and third prizes of $800 and $35, respectively, if exactly five and four of his six selected numbers match the winning numbers.
- What is the expected value of the amount a player wins in one game?

1. Let X be the amount that a player wins in one game. The possible values of X are 1,200,000; 800; 35; and 0. The probabilities are

$$p(X = 1{,}200{,}000) = \frac{1}{\binom{49}{6}} \approx 0.000{,}000{,}072.$$

$$p(X = 800) = \frac{\binom{6}{5}\binom{43}{1}}{\binom{49}{6}} \approx 0.000{,}018. \; p(X = 35) = \frac{\binom{6}{4}\binom{43}{2}}{\binom{49}{6}} \approx 0.000{,}97.$$

$$P(X = 0) = 1 - 0.000{,}000{,}072 - 0.000{,}018 - 0.000{,}97 = 0.999{,}011{,}928.$$

Therefore,

$$E(X) \approx$$

2. This show that on the average players will win 13 cents per game. If the cost per game is 50 cents, then, on the average, a player will lose 37 cents per game. Therefore, a player who plays 10,000 games over several years will loss approximately $

Example 4.20

- The tanks of a country's army are numbered 1 to N.
- In a war this country loses n random tanks to the enemy, who discovers that the captured tanks are numbered.
- If $X_1, X_2, \dots, X_n$ are the numbers of the captured tanks, what is $E(\max X_i)$? How can the enemy use $E(\max X_i)$ to find an estimate of N, the total number of this country's tanks?

1. Let Y = max X_i; then $P(Y=k)=\frac{\binom{k-1}{n-1}}{\binom{N}{n}} \quad k=n, n+1, n+2, \ldots, N,$

$$E(Y)= \qquad = \frac{1}{\binom{N}{n}}\sum_{k=n}^{N}\frac{k(k-1)!}{(n-1)!(k-n)!}$$

$$=\frac{n}{\binom{N}{n}}\sum_{k=n}^{N}\binom{k}{n}$$

2. $\sum_{k=n}^{N}\binom{k}{n}$ is the coefficient of x^n in the polynomial $\sum_{k=n}^{N}(1+x)^k$

Since, $\sum_{k=n}^{N}(1+x)^k = (1+x)^n \sum_{k=0}^{N-n}(1+x)^k = (1+x)^n \frac{(1+x)^{N-n+1}-1}{(1+x)-1}$

$$= \frac{1}{x}\left[(1+x)^{N+1}-(1+x)^n\right]$$

3. The coefficient of x^n in the polynomial is $\binom{N+1}{n+1}$

4. We obtain

$$E(Y) = \qquad\qquad = \frac{n(N+1)}{n+1}$$

5. To estimate N, the total number of this country's tanks, we obtain $N =$

6. For example, the enemy captures 12 tanks and the maximum of the numbers of the tanks captured is 117, then we get $N \approx (13/12) \times 117 - 1 \approx 126$

Theorem 4.1

- *If X is a constant random variable, that is, if $P(X = c) = 1$ for a constant c, then $E(X) =$*
- *Proof:* There is only one possible value for X and that is c; hence $E(X) = c \cdot P(X = c) = c \cdot 1 = c$.

Theorem 4.2

- *Let X be a discrete random variable with set of possible values A and probability mass function p(x), and let g be a real-valued function. Then g(X) is a random variable with*

$$E[g(X)] =$$

Corollary

- *Let X be a discrete random variable; g_1, g_2, . . . , g_n be real-valued functions, and let a_1, a_2, . . . , a_n be real numbers. Then*

$$E[\alpha_1 g_1(X)+\alpha_2 g_2(X)+...+\alpha_n g_n(X)]$$

$$=\alpha_1 E[g_1(X)]+\alpha_2 E[g_2(X)]+...+\alpha_n E[g_n(X)]$$

  - This corollary implies that $E(X)$ is linear. That is, if $\alpha, \beta \in \mathbf{R}$, then
    $\boldsymbol{E(\alpha X + \beta) =}$

$$E(X^2) = \sum_{x \in A} x^2 p(x)$$

$$E(X^2 - 2X + 4) =$$

$$E(X \cos X) = \sum_{x \in A} (x \cos x) p(x)$$

$$E(e^X) = \sum_{x \in A} e^x p(x)$$

Example 4.23

- The probability mass function of a discrete random variable *X* is given by

$$p(x) = \begin{cases} x/15 & x = 1,2,3,4,5 \\ 0 & otherwise \end{cases}$$

What is the expected value of X(6 – X)?

1.

$$E[X(6-X)] =$$

Example 4.24

- A box contains 10 disks of radii 1, 2, . . . , and 10, respectively. What is the expected value of the area of a disk selected at random from this box?

1. Let the radius of the disk be *R*

   R is a random variable with the probability mass function $p(x) = 1/10$ if $x = 1, 2, \ldots, 10$, and $p(x) =$ 0 otherwise.

2. $E(\pi R^2) = \qquad = 38.5\pi$

4.5 Variances and moments of discrete random variables

- **Definition** *Let X be a discrete random variable with a set of possible values A, probability mass function p(x), and E(X) = μ. Then* σ_X *and Var(X), called the* ***standard deviation*** *and the* ***variance*** *of X, respectively, are defined by*

  $\sigma_X = \sqrt{E[(X-\mu)^2]}$ and Var(X) =

  - Note that by this definition and Theorem 4.2,

$$Var(X) = E[(X-\mu)^2] = \sum_{x \in A} (x-\mu)^2 p(x)$$

Example 4.26

- Karen is interested in two games, Keno and Bolita.
- To play Bolita, she buys a ticket for $1, draws a ball at random from a box of 100 balls numbered 1 to 100. If the ball drawn matches the number on her ticket, she wins $75; otherwise, she loses.
- To play Keno, Karen bets $1 on a single number that has a 25% chance to win. If she wins, they will return her dollar plus two dollars more; otherwise, they keep the dollar.
- Let *B* and *K* be the amounts that Karen gains in one play of Bolita and Keno, respectively. Calculate E(B), E(K), Var(B), and Var(K).

1. Let B and K be the amounts that Karen gains in one play of Bolita and Keno, respectively.
2. $E(B) = \qquad\qquad\qquad\qquad\qquad\qquad = -0.25$

   $E(K) = (2)(0.25) + (-1)(0.75) = -0.25$
3. $\mathrm{Var}(B) =$

   $\mathrm{Var}(K) = E[(K - \mu)^2] = (2 + 0.25)^2(0.25) + (-1 + 0.25)^2(0.75) = 1.6875$

Theorem 4.3

- Var(X) = E(X^2) – [E(X)]^2
- *Proof:* By the definition of variance,

$$Var(X) = E[(X-\mu)^2] = E(X^2 - 2\mu X + \mu^2)$$

$$=$$

$$=$$

$$= E(X^2) - [E(X)]^2$$

  - Since Var(X) ≥ 0, for any discrete random variable X,
    [E(X)]^2 ≤ E(X^2)

Example 4.27

- What is the variance of the random variable *X*, the outcome of rolling a fair die?

1. The probability mass function of *X* is given by

   p(x) =

2. $$E(X)=\sum_{x=1}^{6}xp(x)=\frac{1}{6}\sum_{x=1}^{6}x=\frac{1}{6}(1+2+3+4+5+6)=\frac{7}{2}$$

   $$E(X^2)=\sum_{x=1}^{6}x^2p(x)=\frac{1}{6}\sum_{x=1}^{6}x^2=\frac{1}{6}(1+4+9+16+25+36)=\frac{91}{6}$$

   $$\Rightarrow Var(X)=$$

Theorem 4.4

- *Let X be a discrete random variable with the set of possible values A, and mean μ. Then* $\mathrm{Var}(X) = 0$ *if and only if X is a constant with probability*

Theorem 4.5

- *Let X be a discrete random variable; then for constants a and b we have that*

  $\mathrm{Var}(aX + b) =$

  $\sigma_{aX+b} = |a|\sigma_X$

Example 4.28

- Suppose that, for a discrete random variable X, E(X) = 2 and E[X(X − 4)]= 5. Find the variance and the standard deviation of $-4X + 12$.

1. By the Corollary of Theorem 4.2, $E[X^2 - 4X] = 5$ implies that $E(X^2) - 4E(X) = 5$
2. Substituting E(X) in this relation gives $E(X^2) = 13$.
3. $Var(X) = E(X^2) - [E(X)]^2 = 13 - 4 = 9$,
   $\sigma_X = \sqrt{9} = 3$
4. $Var(-4X + 12) =$
   $\sigma_{-4X+12} =$

Definition

- *Let X and Y be two random variables and ω be a given point. If for all t > 0,*

$$P(|Y - \omega| \leq t) \qquad P(|X - \omega| \leq t),$$

then we say that X is more ***concentrated*** *about ω than is Y .*

Theorem 4.6

- *Suppose that X and Y are two random variables with $E(X) = E(Y) = \mu$. If X is more concentrated about μ than is Y, then*

Moment

- Let X be a random variable with expected value μ. Let c be a constant, $n \geq 0$ be an integer, and $r > 0$ be any real number, integral or not. The expected value of X, $E(X)$, is also called the **first moment** of X.

$E[g(X)]$	Definition
$E(X^n)$	The n^{th} moment of X
$E(\|X\|^r)$	The r^{th} absolute moment of X
$E(X-c)$	The 1^{st} moment of X about c
	The n^{th} moment of X about c
$E[(X-\mu)^n]$	The n^{th} central moment of X

4.6 Standardized random variables

- Let X be a random variable with mean μ and standard deviation σ. The random variable $X^* = (X - \mu)/\sigma$ is called the **standardized** X. We have that

$$E(X^*) = E(\frac{1}{\sigma}X - \frac{\mu}{\sigma}) = \frac{1}{\sigma}E(X) - \frac{\mu}{\sigma} =$$

$$Var(X^*) = Var(\frac{1}{\sigma}X - \frac{\mu}{\sigma}) = \frac{1}{\sigma^2}Var(X) =$$

- Standardization is particularly useful if two or more random variables with different distributions must be compared.
- Suppose that, for example, a student's grade in a probability test is 72 and that her grade in a history test is 85.
- Suppose that the mean and standard deviation of all grades in the history test are 82 and 7, respectively, while these quantities in the probability test are 68 and 4.

- If we convert the student’s grades to their standard deviation units, we find that her standard scores on the probability and history tests are given by and , respectively.
- These show that her grade in probability is 1 and in history is 0.43 standard deviation unit higher than their respective averages.
- Therefore, she is doing relatively better in the probability course than in the history course.