Chapter 3 Random Variables and Distributions

Instructor : Dr. Jing Liang
School of Information and Communication Engineering
liangjing@uestc.edu.cn

Outlines

- Random variables and discrete distributions
- Continuous distributions
- Special distributions (Book Chap. 5)
- The cumulative distribution function
- Bivariate distributions
- Marginal distributions
- Conditional distributions
- Multivariate distributions
- Functions of a random variable
- Functions of two or more random variables

Definition of a Random Variable - 1

- **Definition 3.1.1 Random Variable.**

A **real-valued function** defined on a sample space S.

Ex1 (Book Ex3.1.1) Tossing a Coin. Tossing a fair coin ten times. What's the sample space?

S ={HTTTTHTTTH,…}: all sequences of 10 H &T.

Let s denote the sequence (the outcome of the experiment).

Let X denote the real-valued function defined on S that counts the number of heads in each outcome. e.g.,

s =HHHHTTTTTT, $X(s)$=4.

Function X can take the possible values of 0,1,2,…,10.

Y=10- X. Is Y a *R.V.?* Z=2X. Is Z a *R.V.?* Yes!

Definition of a Random Variable - 1

- **Ex2 Measuring a Person's Height.** Randomly choose a student from Dr. Liang's RM class. His or her height X is a ***Random Variable*.**
- **Ex3 (Book Ex 3.1.4)Tossing a Coin**. A fair coin tossed 10 times. Let X be the number of heads in the 10 tosses. X=?

$X=\{x|x=0,1,2,3,4,5,6,7,8,9,10\}$.

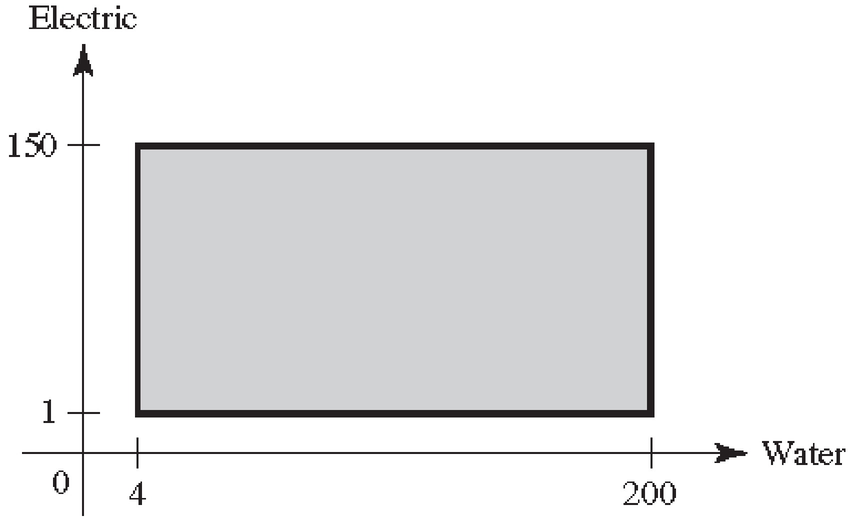
What's $\Pr(X=x)$?

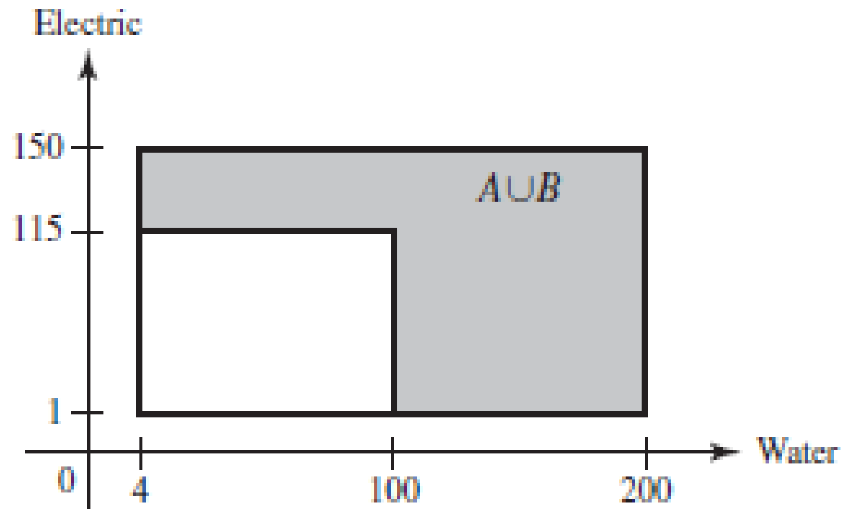
$\Pr(X=x)$ is the sum of the probabilities of all outcomes in the event $\{X=x\}$.

$$\Pr(X=x)=\binom{10}{x}\frac{1}{2^{10}}$$

Definition of a Random Variable - 2

- **Ex4 (Book Ex 3.1.3) Demands for Utilities.**





A is the event that water demand is at least 100 ($100 \leq x \leq 200$). B is the event that electric demand is at least 115 ($115 \leq y \leq 150$).

Define a $R.V.$ $$Z(s) = \begin{cases} 1 & \text{if } s \in A \cup B \\ 0 & \text{if } s \notin A \cup B \end{cases}$$

The Distribution of a *R.V.*- 1

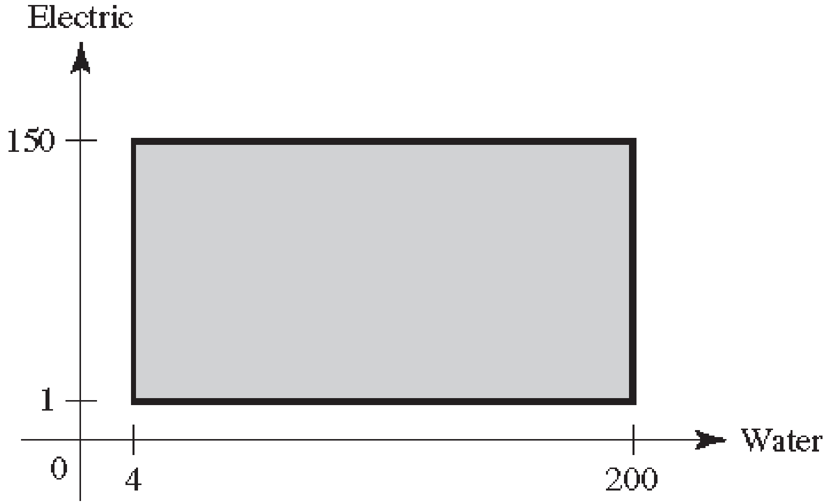
- Let C be a subset of the real line such that $\{X\in C\}$ is an event, and let $\Pr(X\in C)$ denote the prob. that the value of X will belong to the subset C. Then $\Pr(X\in C)$is equal to the prob. that the outcome s of the experiment will be such that $X(s)\in C$: $\Pr(X\in C)=\Pr(\{s: X(s)\in C\})$.
- **Definition 3.1.2 Distribution.**

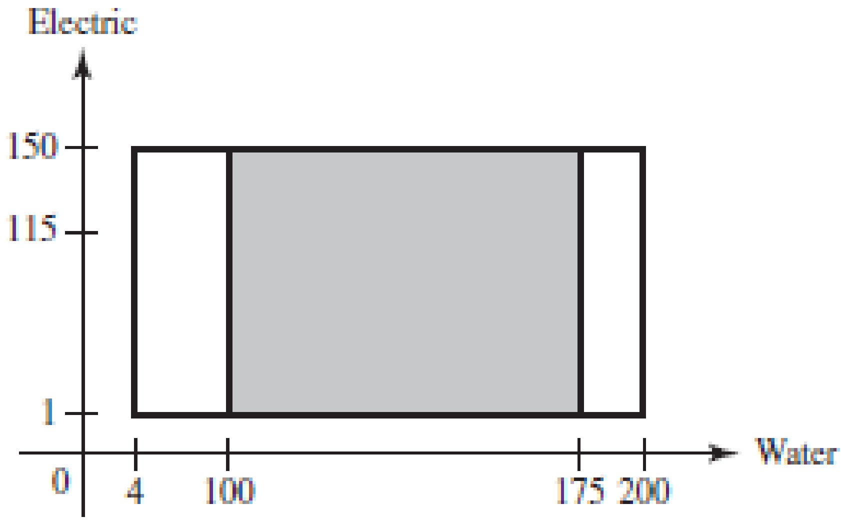
Let X be a *R.V.*. The distribution of X is the **collection of all probabilities** of the form $\Pr(X\in C)$ for **all sets *C* of real numbers** such that $\{X\in C\}$ is an event.

- This distribution is a prob. measure on the set of real numbers. *R.V.* is a main tool used for modeling unknonwn quantities.

The Distribution of a *R.V.*- 2

◆ **Ex5 (Book Ex 3.1.5) Demands for Utilities.**





Let X be the water demand. What's the distribution of X?

$$\Pr(X \in C) = \frac{(150-1)\times(\text{length of interval } C)}{(150-1)\times(200-4)}$$

e.g., C is the interval $[100, 175]$

$$\Pr(X \in C) = 75/196 = 0.3827$$

Discrete Distributions - 1

- **Definition 3.1.3 Discrete Distribution/ Discrete *R.V.***

A *R.V.* X has a ***discrete distribution*** or that X is a ***discrete R.V.*** if it takes a finite number k of different values $x_1, \ldots, x_k$ or, at most, an infinite sequence of different values $x_1, x_2, \ldots$.

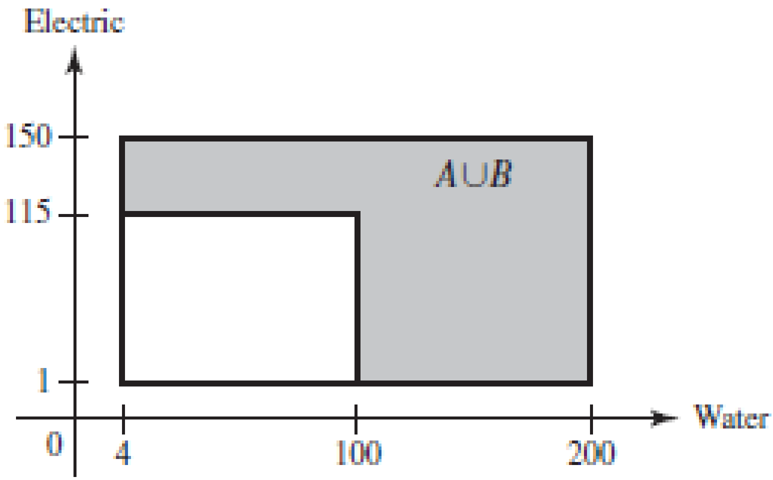
- **Definition 3.1.4 Probability Function/p.f./Support**

If a random variable X has a discrete distribution, the ***probability function (p.f.)*** or ***probability mass function (p.m.f)*** of X is defined as the function f such that for every real number x, $f(x) = \Pr(X = x)$.

The closure of the set $\{x: f(x) > 0\}$ is called ***the support of (the distribution of) X***.

Discrete Distributions - 2

- **Ex6 (Book Ex 3.1.6) Demands for Utilities**



$$Z(s)=\begin{cases}1 & \text{if } s\in A\bigcup B\\ 0 & \text{if } s\notin A\bigcup B\end{cases}$$

If Z has p.f. f, then

$$f(Z)=\begin{cases}0.6525 & \text{if } z=1,\\ 0.3475 & \text{if } z=0,\\ 0 & \text{otherwise.}\end{cases}$$

The support of Z?

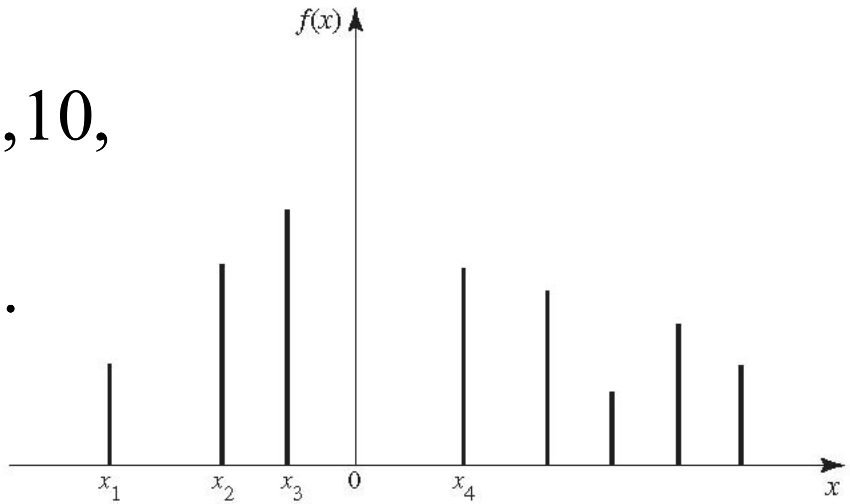
The set {0,1}, only 2 elements.

Discrete Distributions - 3

- **Ex7 (Book Ex 3.1.7) Tossing a Coin.**

A fair coin tossed 10 times. Let X be the number of heads in the 10 tosses. Its p.f. $f(x) = ?$

$$f(x) = \begin{cases} \binom{10}{x} \dfrac{1}{2^{10}} & x = 0, 1, \cdots, 10, \\ 0 & \text{otherwise.} \end{cases}$$



An ex. of a p.f.

What's the support of X?

The set $\{0,1,\ldots,10\}$.

The sum of the heights of the vertical segments in the above figure is ?

Simple facts about p.f.

Let X be a discrete $R.V.$ (X has a discrete distribution) with p.f. f.

$f(x) \geq 0$ If x is not the possible values of X, = holds.

$$\sum_{i=1}^{\infty} f(x_i) = 1$$ **Theorem 3.1.1**

$$\Pr(X \in C) = \sum_{x_i \in C} f(x_i)$$ **Theorem 3.1.2**

Ex8 Suppose that a $R.V.$ X has a discrete distribution with the following p.f.:

$$f(x) = \begin{cases} cx & \text{for } x = 1, \cdots, 5, \\ 0 & \text{otherwise.} \end{cases}$$

What's the value of c?

Sol: $\because \sum_{i=1}^{5} f(x_i) = 1$

$$c\sum_{i=1}^{5} i = 1 \Rightarrow c = \frac{1}{15}$$

Bernoulli Distribution / *R.V.*

- **Definition 3.1.5** A *R.V. X* that takes **only two values** 0 and 1 with Pr $(X=1)=p$ has the ***Bernoulli distribution with parameter p***, *or X* is a ***Bernoulli R.V. with parameter p.*** The p.f. of *X* can be written as follows:

$$\boxed{f(x)=\begin{cases} p^x(1-p)^{1-x} & \text{for } x=0,1, \\ 0 & \text{otherwise.} \end{cases}}$$

- **Definition 5.2.2** **Bernoulli Trials/Process.**

If the *R.V.* in a finite or infinite sequence $X_1, X_2, \ldots$ are ***independent and identically distributed (i.i.d.)***, and if each *R.V.* X_i has the Bernoulli distribution with parameter p, then $X_1, X_2, \ldots$ are ***Bernoulli trials*** with parameter p. An **infinite** sequence of Bernoulli trials is also called a ***Bernoulli process***.

Bernoulli Trials/Process Ex.

- **Ex9 (Book Ex 5.2.2)Tossing a Coin**. Suppose that a fair coin is tossed repeatedly. Let $X_i = 1$ if a head is obtained on the i^{th} toss, and let $X_i = 0$ if a tail is obtained ($i = 1, 2, \ldots$). Then the random variables $X_1, X_2, \ldots$ are Bernoulli trials with parameter $p = 1/2$.
- **Ex10 (Book Ex 5.2.3) Defective Parts.** Suppose that 10 percent of the items produced by a certain machine are defective and the items are independent of each other. We will sample n items at random and inspect them. Let $X_i = 1$ if the i^{th} item is defective, and let $X_i = 0$ if it is nondefective ($i = 1, \ldots, n$). Then the variables $X_1, \ldots, X_n$ form n Bernoulli trials with parameter $p = 1/10$.

Binomial Distributions - 1

- **Ex11 (Book Ex3.1.9) Defective Parts.** Suppose a machine produces a defective item with prob. p and a nondefective item with prob. 1-p. Examine n items. Let X denote the number of items that are defective.

$$\Pr(X = x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

The p.f. of X will be

$$\boxed{f(x) = \begin{cases} \boxed{\binom{n}{x} p^x (1-p)^{n-x}} & \text{for } x = 0,1,2,\ldots,n, \\ 0 & \text{otherwise.} \end{cases}}$$

Ex4 is a Binomial *R.V.*

Definition 3.1.7 Binomial Distribution

A binomial *R.V.* with parameters *n* and *p*.

$$b(x; n, p)$$

Binomial Distributions - 2

- **Ex12** Suppose that X is a Binomial $R.V.$ with parameters n=15 and p=0.5. Find Pr (X < 6).

$$\Pr(X < 6) = \sum_{k=0}^{5} f(x = k)$$

Table of Binomial Probabilities in Book page 790.

n	k	$p = 0.1$	$p = 0.2$	$p = 0.3$	$p = 0.4$	$p = 0.5$
15	0	.2059	.0352	.0047	.0005	.0000
	1	.3432	.1319	.0305	.0047	.0005
	2	.2669	.2309	.0916	.0219	.0032
	3	.1285	.2501	.1700	.0634	.0139
	4	.0428	.1876	.2186	.1268	.0417
	5	.0105	.1032	.2061	.1859	.0916

Sum them up

$$\Pr(X < 6) = 0.1509$$

Q1: Is the binomial p.f. symmetric about $n/2$?

Q2: How to obtain the p.f. with p>0.5 and p=0.55?

Binomial Distributions - 3

$$\Pr(X = n-k \mid n, 1-p) = \Pr(X = k \mid n, p)$$

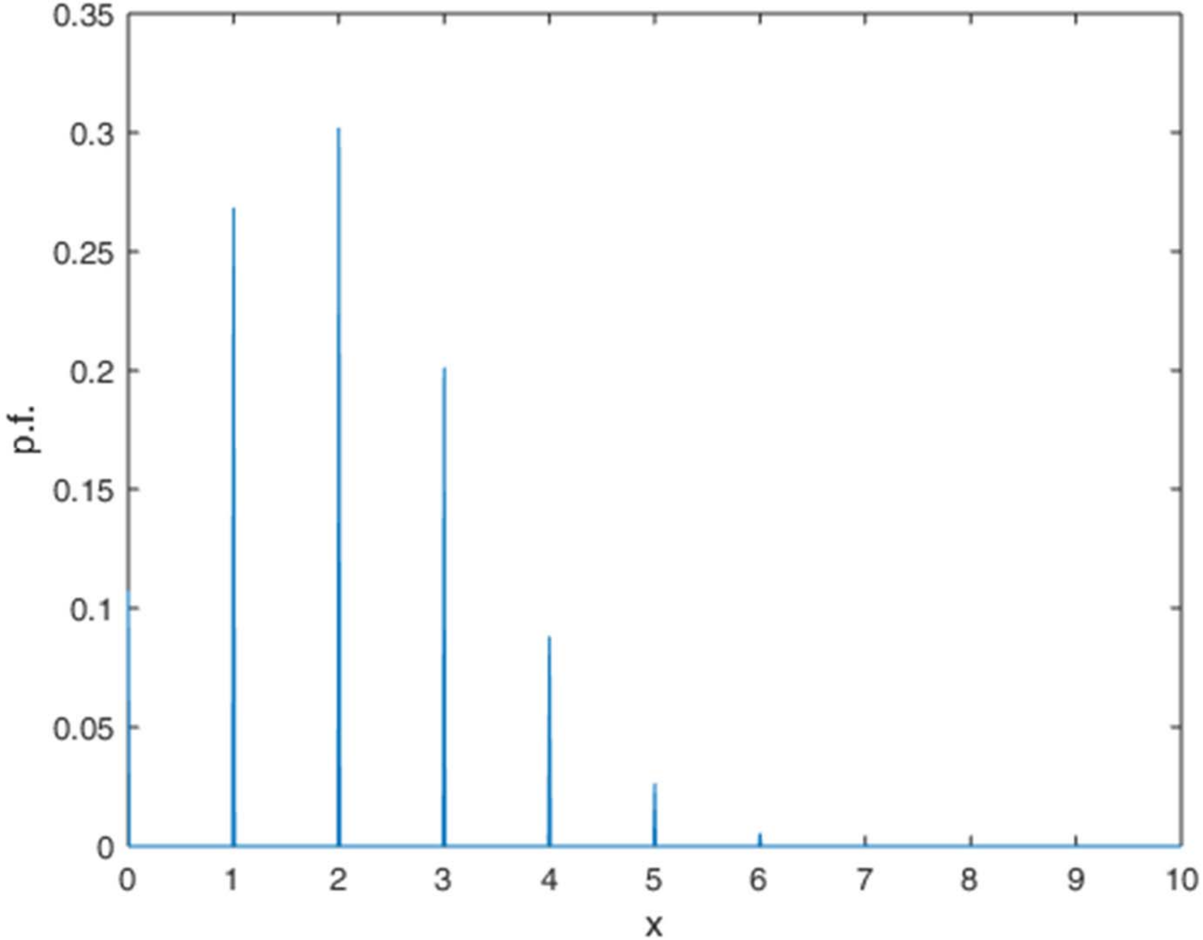


Figure 1 The binomial distribution(n=10,p=0.2)

Binomial Distributions - 4

- Where does the name binomial come from?

From the fact that

$$(q+p)^n = \sum_{k=0}^{n} \binom{n}{k} q^k p^{n-k} = \sum_{k=0}^{n} b(k;n,p)$$

If q=1-p, we have p+q=1, and thus

$$\sum_{k=0}^{n} b(k;n,p) = 1.$$

Binomial and Bernoulli

- **Theorem 5.2.1** If the *R.V.*s $X_1, \ldots, X_n$ form n Bernoulli trials with parameter p, and if $X = X_1 + \ldots + X_n$, then X has the binomial distribution with parameters n and p.

Theorem 5.2.2 If $X_1, \ldots, X_k$ are independent *R.V.s*, and if X_i has the binomial distribution with parameters p and n_i ($i=1,\ldots,k$), then the sum $X_1 + \ldots + X_k$ has the binomial distribution with parameters $n = n_1 + \ldots + n_k$ and p.

Multinomial Distribution-1

- If a given trial can result in partition of k *disjoint* possible events, $X_1, \ldots, X_k$ with prob. $p_1, p_2, \ldots, p_k$, then the multinomial distribution will give the prob. that X_1 occurs x_1 times, X_2 occurs x_2 times ,…, and X_k occurs x_k times in n independent trials, where $x_1 + x_2 + \ldots + x_k = n$ *and* $p_1 + p_2 + \ldots + p_k = 1$. What's the joint distribution $f(\underline{x})$?

$$f(\underline{x}) = \begin{cases} \binom{n}{x_1, x_2, \cdots, x_k} p_1^{x_1} p_2^{x_2} \cdots p_k^{x_k} & \text{if } x_1 + x_2 + \cdots + x_k = n, \\ 0 & \text{otherwise.} \end{cases}$$

Definition 5.9.1 Multinomial Distributions A discrete $R.V.\ \underline{X} = (X_1, \ldots, X_k)$ whose p.f. is shown above has the multinomial distribution with parameters n and $\underline{p} = (p_1, p_2, \ldots, p_k)$.

Multinomial Distribution-2

◆ **Ex13 Select balls (Ex 5.9.2)** Suppose the probabilities to randomly select a red, a blue and a white ball are 0.23, 0.59, 0.18, respectively. All balls are distinctive only in colour. Now 20 balls have been randomly selected with the replacement. Determine the probability that 7 are red, 8 are blue and 5 are white.

Sol:

$$\frac{20!}{7!8!5!}\times 0.23^{7}\times 0.59^{8}\times 0.18^{5}=0.0094$$

Uniform Distributions on Integers-1

Ex14 (Ex 3.1.8) Daily Numbers. A lottery game requires to select a three-digit number (leading 0s allowed). The sample space here consists of all triples (i_1,i_2,i_3) where $i_j \in \{0,\ldots,9\}$ for j=1,2,3. If s = (i_1,i_2,i_3), define $X(s) = 100i_1+10i_2+i_3$.

e.g., $X(0,1,5)= 15$.

$\Pr(X=x) = 0.001$ for each integer $x \in \{0,1,\ldots,999\}$.

The X in Ex14 has the ***uniform distribution*** on the integers 0,1,…,999.

A uniform distribution on a set of k integers has prob. $1/k$ on each integer, or we say that one of the k integers are chosen **at random.**

Uniform Distributions on Integers-2

- **Definition 3.1.6** **Uniform Distribution on Integers.**

Let a, b ($a \le b$) be integers. Suppose that the value of a *R.V. X* is equally likely to be each of the integers $a,\ldots,b$. Then we say that X has the ***uniform distribution on the integers a,…,b.***

If $b > a$, there are $b- a+1$ integers from a to b including a and b.

- **Theorem 3.1.3** If X has the uniform distribution on the integers $a,\ldots,b,$ the p.f. of X is

$$f(x) = \begin{cases} \dfrac{1}{b-a+1} & \text{for } x = a,\ldots,b, \\ 0 & \text{otherwise.} \end{cases}$$

Uniform Distributions on Integers-3

- **Ex15** Suppose that a $R.V.$ X has the uniform distribution on the integers 10,…,20. Find the probability that X is even.

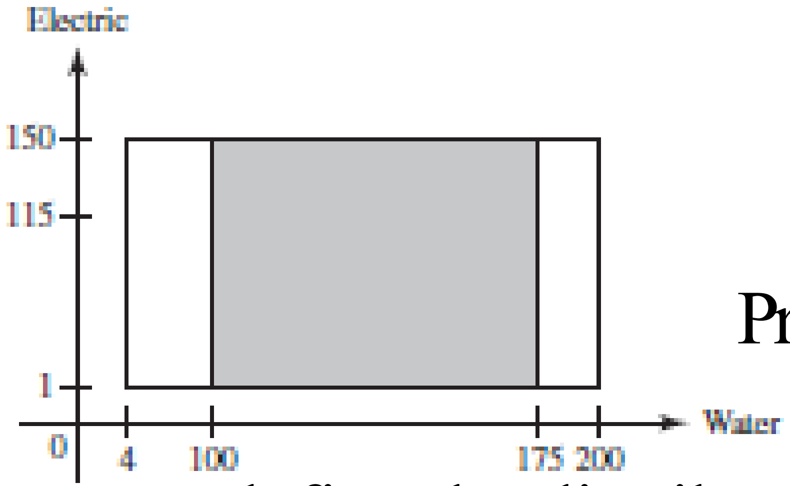
Sol: we have

$$f(x) = \begin{cases} \dfrac{1}{11} & \text{for } x = 10,\ldots,20, \\ 0 & \text{otherwise.} \end{cases}$$

Since there are 6 even numbers from 10 to 20, the probability that X is even should be 6/11.

Continuous Distributions - 1

◆ **Ex16 (Ex 3.2.1) Demands for Utilities.** Determine the distribution of the demand for water X.



Sol: For each interval

$$C=[C_0,C_1]\subset[4,200]$$

$$\Pr(c_0\le X\le c_1)=\frac{c_1-c_0}{196}=\int_{c_0}^{c_1}\frac{1}{196}dx$$

We can define the distribution

$$f(x)=\begin{cases}\frac{1}{196} & \text{if } 4\le x\le 200,\\ 0 & \text{otherwise.}\end{cases}$$

$$\Pr(c_0\le X\le c_1)=\int_{c_0}^{c_1}f(x)dx$$

The above equation holds even if $c_0=-\infty$ or $c_1=\infty$.

Continuous Distributions - 2

- **Definition 3.2.1 Continuous Distribution/R.V.**

A $R.V.$ X has a ***continuous distribution*** or that X is a ***continuous R.V.*** if there exists a **nonnegative function** f, defined on the **real** line, such that for every interval of **real numbers** (bounded or unbounded), the probability that X takes a value in the interval is the **integral** of f over the interval.

e.g., $\Pr(a \le X \le b) = \int_a^b f(x)dx, \quad \Pr(X \le b) = \int_{-\infty}^b f(x)dx,$

$$\Pr(X \ge a) = \int_a^{\infty} f(x)dx.$$

Here $f(x)$ or f is similar to p.f. for discrete $R.V.$.

Continuous Distributions - 3

- **Definition 3.2.2 Probability Density Function/p.d.f.**

If X has a ***continuous distribution***, the function ***f*** described in above Definition 3.2.1 is the ***probability density function*** (abbreviated **p.d.f.**) of X. The closure of the set $\{x : f(x) > 0\}$ is called ***the support of*** (the distribution of) X.

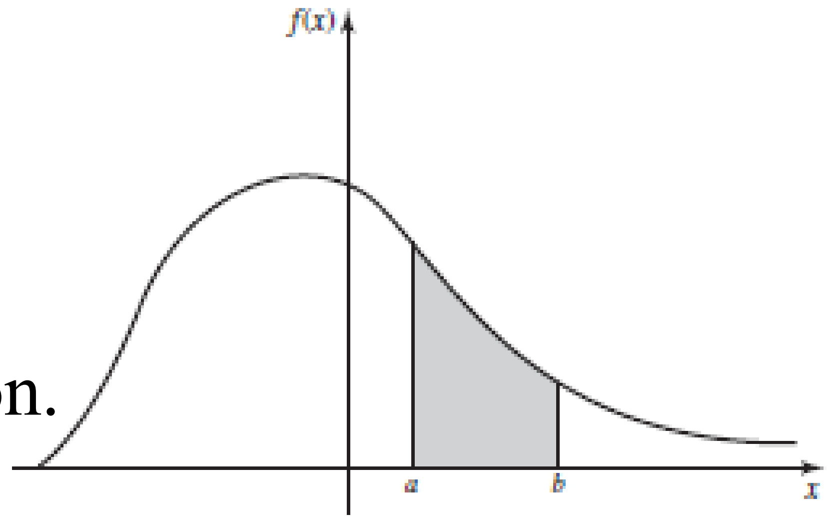
- Every p.d.f. must satisfy two requirements:

$$f(x) \geq 0, \text{for all } x,$$

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

$$\Pr(a \leq X \leq b) = ?$$

The area of the shaded region.



Continuous Distributions - 4

- Note: continuous distributions assign probability 0 to individual values.
- If X has a continuous distribution,

$$\Pr(a \le X \le b) = \Pr(a \le X < b) = \Pr(a < X \le b) = \Pr(a < X < b)$$

$\Pr(X=b) = 0$ for each number b.

$\Pr(X=b)=0$ does not imply that $X=b$ is impossible.

$X=b$ is possible even if we have $\Pr(X=b)=0$.

If $\Pr(X=b)=0$ meant $X=b$ is impossible, all values of X would be impossible and X couldn't assume any value.

The prob. distribution of X is spread so thinly!

The same as the fact that lines have 0 area in two dimensions, but does not mean that no lines.

Continuous Distributions - 5

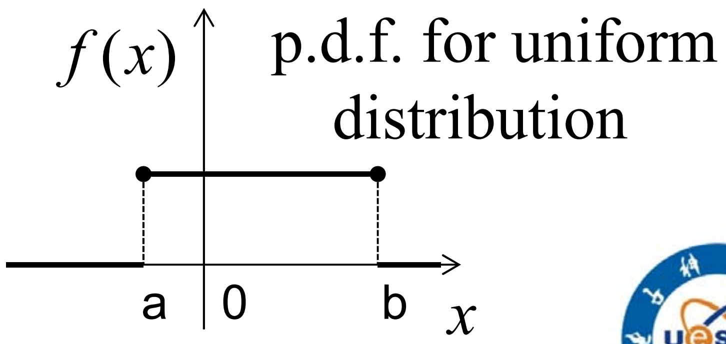
- Due to the property that $\Pr(X=x)=0$ for every individual value x, p.d.f. can be changed at a finite number of points, the p.d.f. of a *R.V.* is not unique.
- However, in this class we adopt the following practice: If a *R.V.* has continuous distribution, we shall give only one version of the p.d.f. of X, just as thought it had been uniquely determined.
- The support of a continuous distribution is the closure of the set where the p.d.f. is strictly positive. It can be shown that the support is unique.

Uniform Distributions on Intervals

Definition 3.2.3 Let a and b be two given real numbers such that $a < b$. Let X be a $R.V.$ such that it is known that $a \le X \le b$ and, for every subinterval of $[a, b]$, the probability that X will belong to that subinterval is proportional to the length of that subinterval. The $R.V. X$ has the ***uniform distribution on the interval*** $[a, b]$.

Theorem 3.2.1 Uniform Distribution p.d.f. If X has the uniform distribution on an interval $[a,b]$, then the p.d.f of X is

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{if } a \le x \le b, \\ 0 & \text{otherwise.} \end{cases}$$



Unbounded *R.V.s*

- **Ex18 (Ex 3.2.5)** The voltage X in a certain electrical system might be a $R.V.$ with a continuous distribution that can be approximately represented by the following p.d.f. What's Pr $(X \le 4)$?

$$f(x) = \begin{cases} 0 & \text{for } x \le 0, \\ \dfrac{1}{(1+x)^2} & \text{for } \boxed{x>0.} \end{cases}$$

unbounded interval

It satisfies both:

$$f(x) \ge 0, \text{for all } x,$$

$$\int_{-\infty}^{\infty} f(x)dx = 1.$$

$$\Pr(X \le 4) = \int_{-\infty}^{0} f(x)dx + \int_{0}^{4} f(x)dx$$

$$= \left.\frac{-1}{1+x}\right|_0^4$$

$$= \frac{4}{5}$$

$$\Pr(X > 1000) \approx 0.001$$

Unbounded p.d.f.'s.

- A value of a **p.d.f.** is a probability density, rather than a probability, such a value **can be larger than 1**.

Ex19 (Book Ex 3.2.6)

$$f(x)=\begin{cases}\dfrac{2}{3}x^{-1/3} & \text{for } 0< x<1,\\ 0 & \text{otherwise.}\end{cases}$$

The values of the p.d.f. are unbounded in the neighborhood of x=0.

- **Density ≠ Probability**
- The values of p.d.f. can be greater than 1, probability is never greater than 1.
- The values of p.d.f. can be unbounded, probability is bounded.

The Poisson Distributions - 1

- **Definition 5.4.1** Let λ>0, a *R.V. X* has the ***Poisson distribution with mean λ*** if the p.f. of *X* is as follows

$$f(x)=\begin{cases}\dfrac{e^{-\lambda}\lambda^{x}}{x!} & x=0,1,2,\cdots,\\ 0 & otherwise.\end{cases}$$

- Poisson *R.V. X* is used to model the number of arrivals that occur in a fixed time period.
- E.g., customers arrive at KFC at a rate of 4.5 customers per hour on average.
- λ can also represent the rate of occurrence of distance, area, volume, etc..

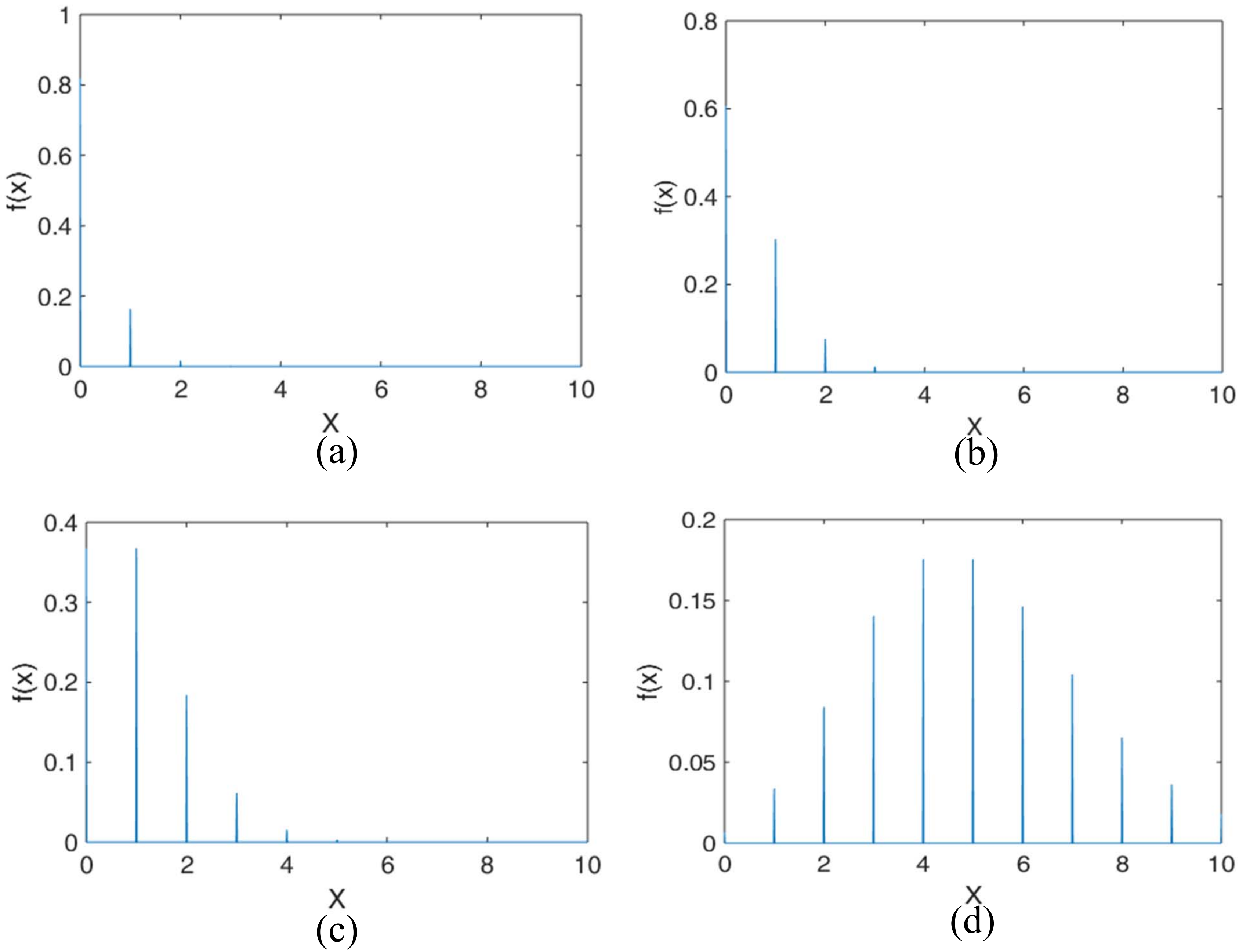


Fig.2 The p.f. of Poisson *R.V*: (a) λ=0.2; (b) λ=0.5; (c) λ=1; (d) λ=5.

The Poisson Distributions - 2

- **Ex20** Suppose that the number of accidents at a certain intersection in Chengdu has the Poisson distribution with mean 0.7 on a given weekend. What's the prob. that there will be at least 3 accidents at the intersection during the weekend?

Sol: from the Table of Poisson Probabilities

k	$\lambda = .1$	.2	.3	.4	.5	.6	.7	.8	.9	1.0
0	.9048	.8187	.7408	.6703	.6065	.5488	.4966	.4493	.4066	.3679
1	.0905	.1637	.2222	.2681	.3033	.3293	.3476	.3595	.3659	.3679
2	.0045	.0164	.0333	.0536	.0758	.0988	.1217	.1438	.1647	.1839
3	.0002	.0011	.0033	.0072	.0126	.0198	.0284	.0383	.0494	.0613
4	.0000	.0001	.0003	.0007	.0016	.0030	.0050	.0077	.0111	.0153
5	.0000	.0000	.0000	.0001	.0002	.0004	.0007	.0012	.0020	.0031
6	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0002	.0003	.0005
7	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
8	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000

Sum them up

The Poisson Distributions - 3

- **Ex21** On the average a store serves 15 customers per hour. What's the prob. that the store will serve more than 20 customers in a particular two-hour period?

Sol: assume that the number of customers served in two-hour period is a Poisson $R.V.$.

$$f(x) = \begin{cases} \dfrac{e^{-\lambda}\lambda^x}{x!} & x = 0,1,2,\cdots, \\ 0 & otherwise. \end{cases}$$

In this case, λ=15×2=30.

$$\Pr(X > 20) = \sum_{x=21}^{\infty} \frac{e^{-30} 30^x}{x!}$$

The Poisson Distributions - 4

- **Theorem 5.4.4** If the *R.V.*s $X_1, \ldots, X_k$ are independent and if X_i has the Poisson distribution with mean λ_i ($i=1,\ldots k$), then the sum $X_1 + \ldots + X_k$ has the Poisson distribution with mean $\lambda_1 + \ldots + \lambda_k$.
- **Theorem 5.4.5 Closeness of Binomial and Poisson.**

For each integer n and each $0<p<1$,let $f(x|n, p)$ denote the p.f. of the binomial distribution with parameters n and p. Let $f(x|\lambda)$ denote the p.f. of the Poisson distribution with mean λ. Let $\{p_n\}_{n=1}^{\infty}$ be a sequence of numbers between 0 and 1 such that $\lim_{n\to\infty} np_n = \lambda$. Then

$$\lim_{n\to\infty} f(x\,|\,n, p_n) = f(x\,|\,\lambda),\ \text{for all } x=0,1,\ldots.$$

If n is large and p is small so that np is close to λ, then the binomial is close to the Poisson distribution.

The Poisson Distributions - 5

- **Ex22** Suppose that the proportion of colorblind people in a certain population is 0.005. What's the prob. that there will not be more than one colorblind person in a randomly chosen group of 600 people?

Q: the number of colorblind person in a randomly chosen group of 600 people is a *R.V. X.* What's the distribution of *X*?

Sol: Binomial.

600 is large and 0.005 is small. It can be approximated by a Poisson distribution with mean λ =600×0.005=3.

It is found from Poisson Table that

$\Pr(X \le 1) = 0.0498 + 0.1494 = 0.1992.$

The Normal Distributions - 1

- **Definition 5.6.1** A $R.V.\ X$ has the ***normal distribution*** with mean μ and variance σ^2 ($-\infty<\mu<\infty$ and $\sigma>0$) if X has a continuous distribution with the following p.d.f.:

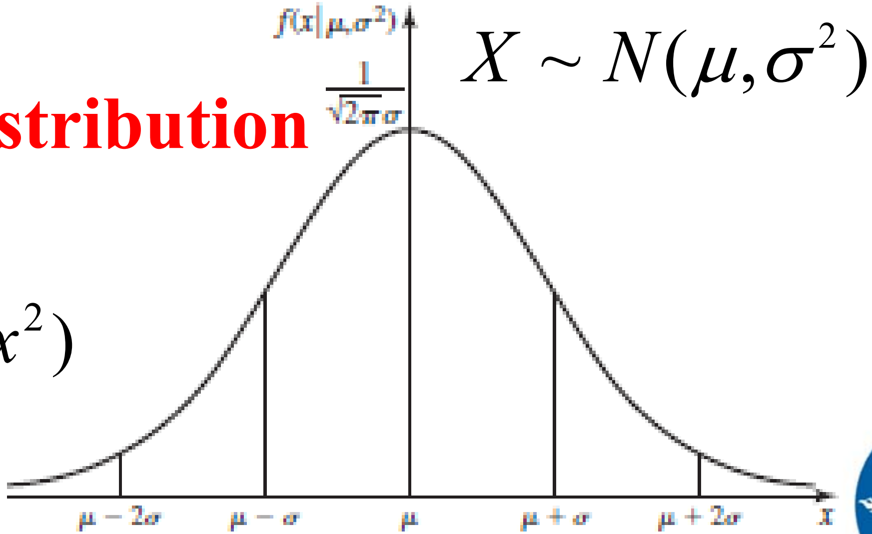
$$\boxed{f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \quad \text{for } -\infty<x<\infty}$$

- **Definition 5.6.2**
Standard Normal Distribution

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

$$\phi(x)=\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{1}{2}x^2\right)$$

for $-\infty<x<\infty$



The Normal Distributions - 2

$$\Phi(x)=\int_{-\infty}^{x}\phi(u)du=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}u^2)du \quad \text{for } -\infty < x < \infty$$

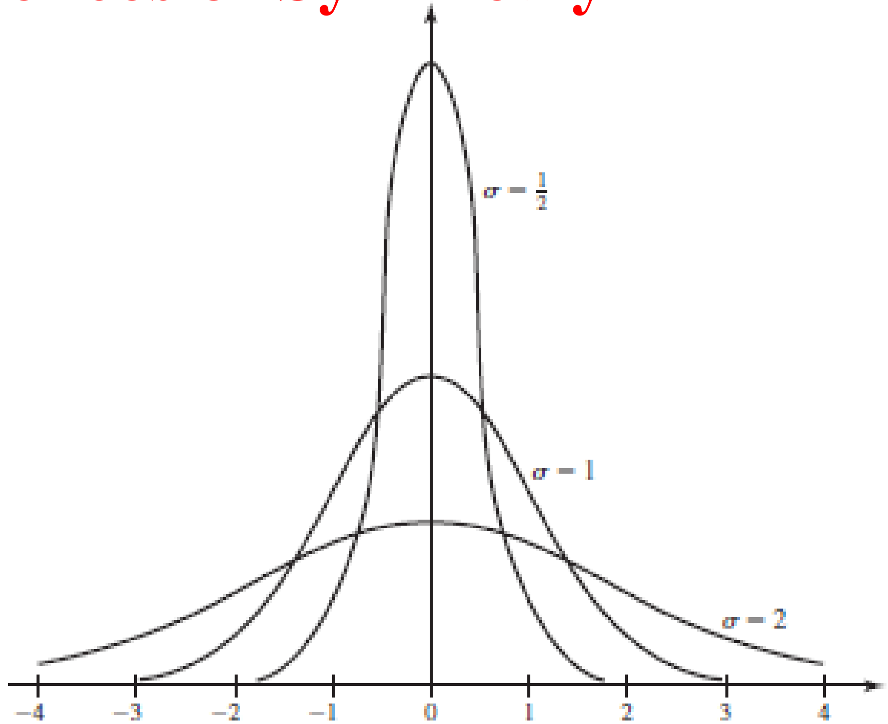
$\Phi(x)$ can not be expressed in closed form in terms of elementary functions. Approximated or by Table.

- **Theorem 5.6.5 Consequences of Symmetry**

For all x and all $0<p<1$,

$$\Phi(-x)=1-\Phi(x)$$

$$\Phi^{-1}(p)=-\Phi^{-1}(1-p)$$



The Normal Distributions - 3

- **Theorem 5.6.6 Converting Normal Distribution to Standard.** Let X have the normal distribution with mean μ and variance σ^2. Then $Z=(X-\mu)/\sigma$ has the standard normal distribution, and for all x

$$\Pr(X \le x) = \Pr(Z \le \frac{x-\mu}{\sigma}) = \Phi(\frac{x-\mu}{\sigma})$$

- **Ex23 (Book Ex5.6.4) Determine probabilities.** Suppose $X \sim N(5, 4)$. Determine the value of $\Pr(1<X<8)$.

$$\Pr(1 < X < 8) = \Pr(\frac{1-5}{2} < \frac{X-5}{2} < \frac{8-5}{2}) = \Pr(-2 < Z < 1.5)$$

$$\Pr(-2 < Z < 1.5) = \Pr(Z < 1.5) - \Pr(Z \le -2)$$

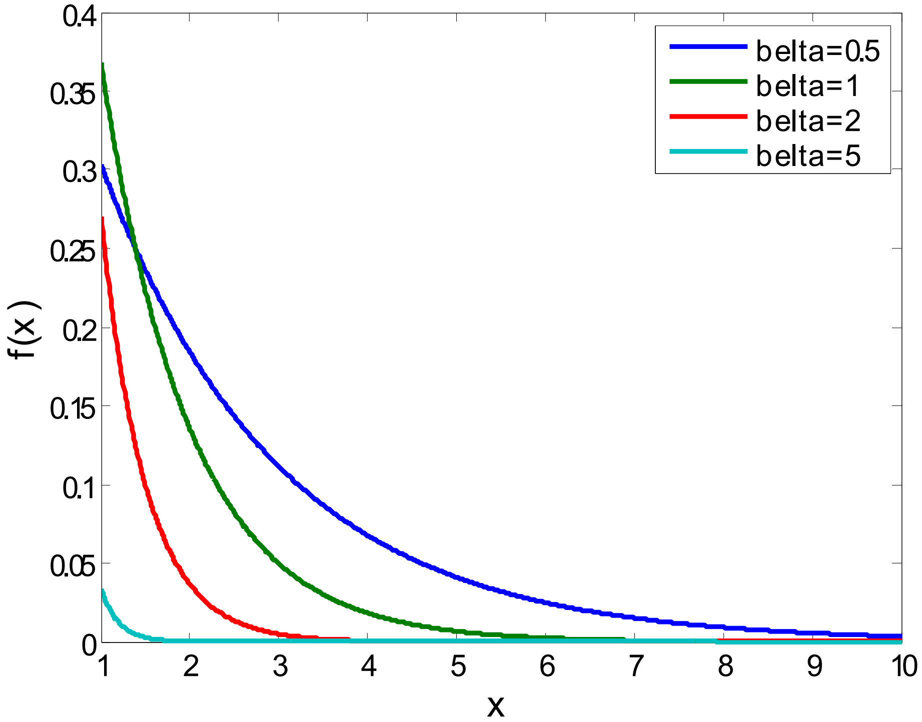
$$= \Phi(1.5) - \Phi(-2)$$

$$= \Phi(1.5) - [1 - \Phi(2)]$$

The Exponential Distribution - 1

◆ **Definition 5.7.3** A $R.V.$ X has the exponential distribution with parameter β ($\beta > 0$) if X has a continuous distribution with the following p.d.f.:

$$f(x) = \begin{cases} \beta e^{-\beta x} & \text{for } x > 0, \\ 0 & \text{for } x \le 0. \end{cases}$$



The Exponential Distribution - 2

- **Ex24** Suppose that a system contains a certain type of component whose time to failure is given by *T* years. The *R.V. T* is modeled nicely by the exponential distribution with β=0.2. If 5 of these components are installed in different systems, what's the prob. that at least 2 are still functioning at the end of 8 years?

Sol: the prob. that a given component is functioning after 8 years is:

$$\Pr(T > 8) = \frac{1}{5}\int_8^{\infty} e^{-t/5} dt = e^{-8/5} \approx 0.2.$$

Let *X* represent the number of components functioning after 8 years. Then *X* is a Binomial *R.V.*.

$$\Pr(X \geq 2) = 1 - \sum_{x=0}^{1} \binom{5}{x} 0.2^x 0.8^{5-x} = 0.2627.$$

Cumulative Distribution Function-1

- **Ex25 (Book Ex 3.3.1)** Voltage X with p.d.f.:

$$f(x)=\begin{cases}0 & \text{for } x\le 0,\\ \dfrac{1}{(1+x)^2} & \text{for } x>0.\end{cases}$$

An alternative characterization that is more directly related to probabilities associated with X:

$$F(x)=\Pr(X\le x)=\int_{-\infty}^{x} f(y)dy=\begin{cases}0 & \text{for } x\le 0,\\ \displaystyle\int_0^x \frac{dy}{(1+y)^2} & \text{for } x>0.\end{cases}$$

$$=\begin{cases}0 & \text{for } x\le 0,\\ 1\text{-}\dfrac{1}{1+x} & \text{for } x>0.\end{cases}$$

$$\Pr(X\le 3)=F(3)=3/4.$$

Cumulative Distribution Function-2

Definition 3.3.1 The ***distribution function*** or ***cumulative distribution function*** (abbreviated c.d.f.) F of a $R.V. X$ is the function

$$F(x) = \Pr(X \leq x) \quad \text{for } -\infty < x < \infty$$

- c.d.f. is defined as above for every $R.V.\ X$.
- c.d.f. is regardless of whether the distribution of X is discrete, continuous or mixed.

Cumulative Distribution Function-3

- **Ex26 (Book Ex 3.3.2) Bernoulli c.d.f**

Let X have the Bernoulli distribution with parameter p, Then the c.d.f. of X is

$$F(x)=\begin{cases}0 & \text{for } x<0,\\ 1-p & \text{for } 0\le x<1\\ 1 & \text{for } x\ge 1.\end{cases}$$

$F(x)$ is the probability of the event $\{X\le x\}$.

$$0\le F(x)\le 1.$$

Properties of c.d.f. - 1

- **Property 3.3.1 Nondecreasing.**

  If $x_1 < x_2$, then $F(x_1) \le F(x_2)$.

Proof: for $x_1 < x_2, \{X \le x_1\} \subset \{X \le x_2\} \to \Pr(X \le x_1) \le \Pr(X \le x_2)$

- **Property 3.3.2 Limits at $\pm\infty$.**

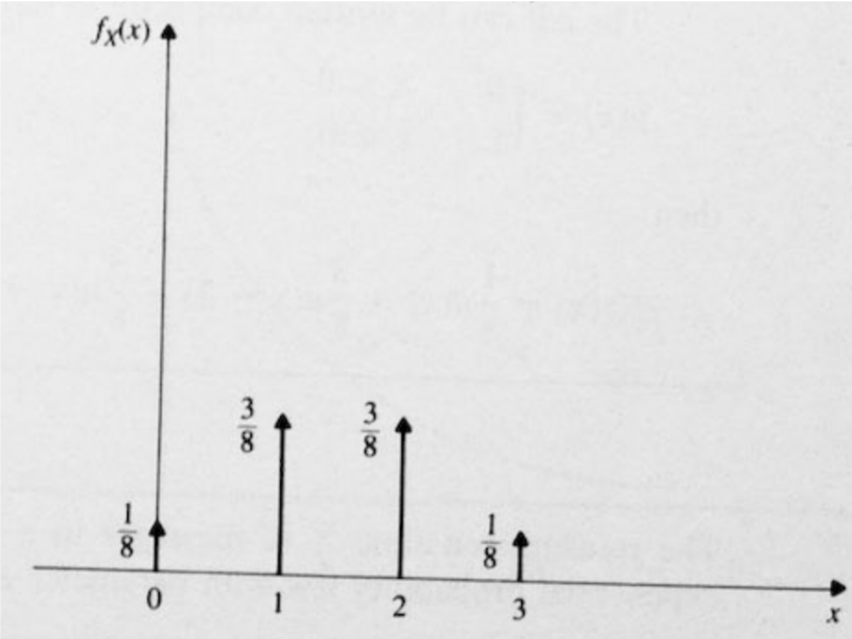
$$\lim_{x\to-\infty} F(x) = \Pr(X \le -\infty) = 0, \lim_{x\to\infty} F(x) = \Pr(X \le \infty) = 1$$

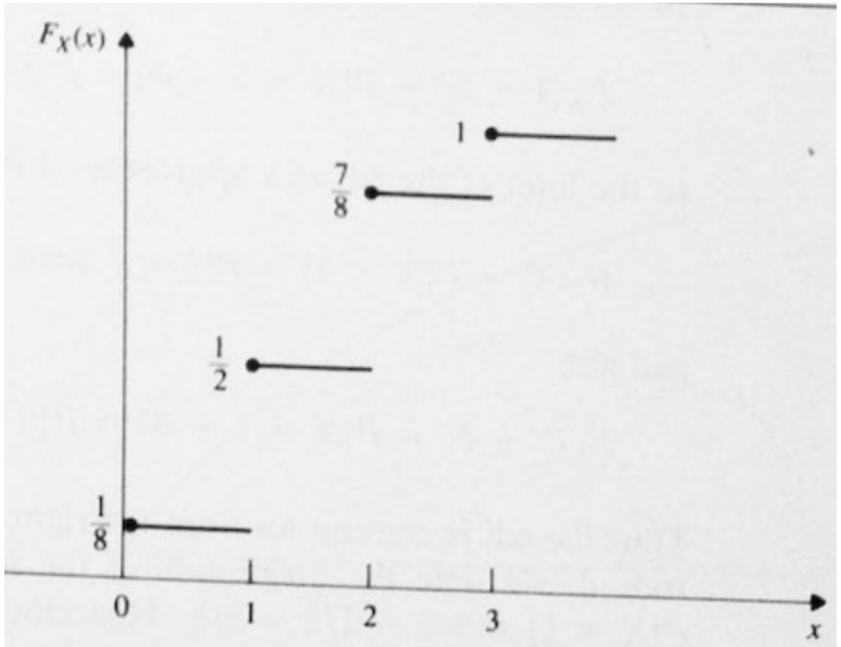
denote $F(x^-) = \lim\limits_{\substack{y\to x\\ y<x}} F(y)$, $F(x^+) = \lim\limits_{\substack{y\to x\\ y>x}} F(y)$, if the c.d.f is continous at a given point x, then $F(x^-) = F(x^+) = F(x)$.

- **Property 3.3.3 Continuity from the Right.**

  $F(x) = F(x^+)$ at every point x.

Properties of c.d.f. - 2

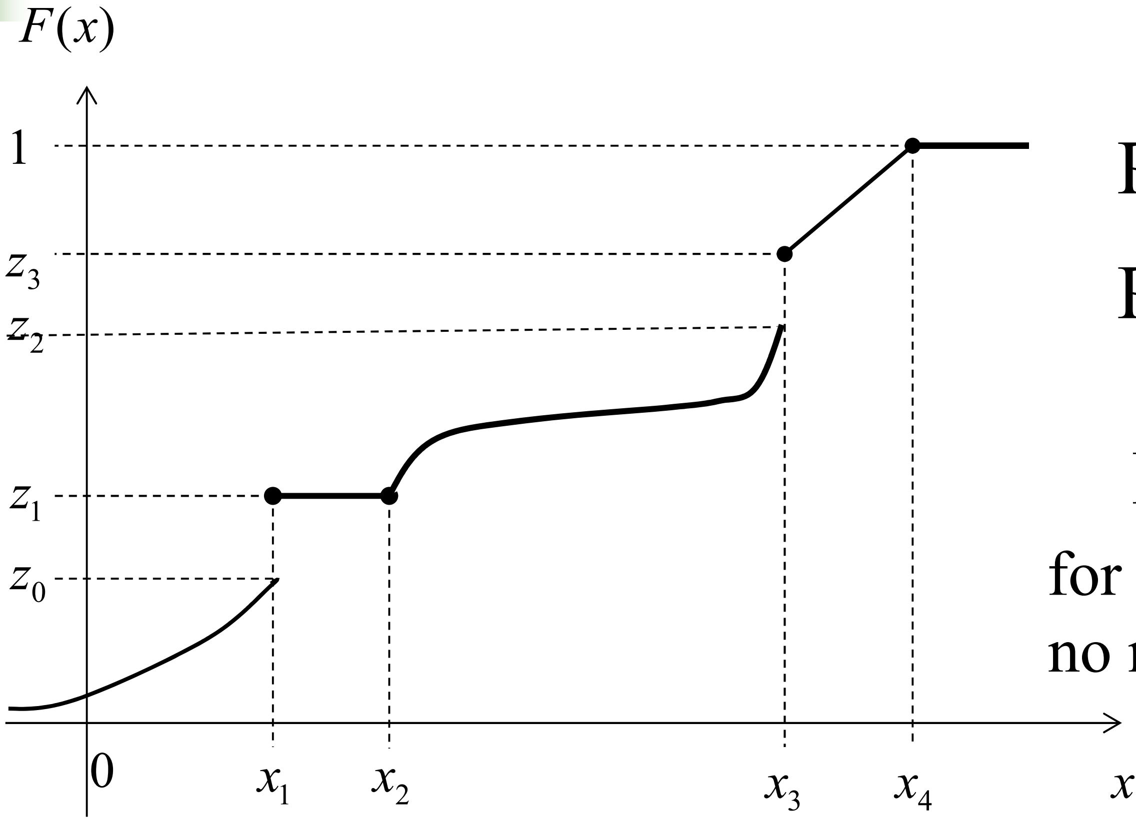




An example to show that

$F(x) = F(x^+)$ at every point x.

Properties of c.d.f. - 3



$$\Pr(X \leq x_4) = 1$$

$$\Pr(X > x_4) = 0$$

$$\Pr(X \leq x) > 0$$

for every finite value of x no matter how small it is.

An example of a c.d.f.

Properties of c.d.f. - 4

- **Theorem 3.3.1** For every value x

$$\Pr(X > x) = 1\text{-}F(x).$$

Proof hint: event $\{X>x\}$ and $\{X\le x\}$ form the partition of S.

- **Theorem 3.3.2** For all values x_1 & x_2 such that $x_1 < x_2$,

$$\Pr(x_1 < X \le x_2) = F(x_2) - F(x_1).$$

Proof hint: $\Pr(x_1 < X \le x_2) + \Pr(X \le x_1) = \Pr(X \le x_2)$

- **Theorem 3.3.3** For each value x,

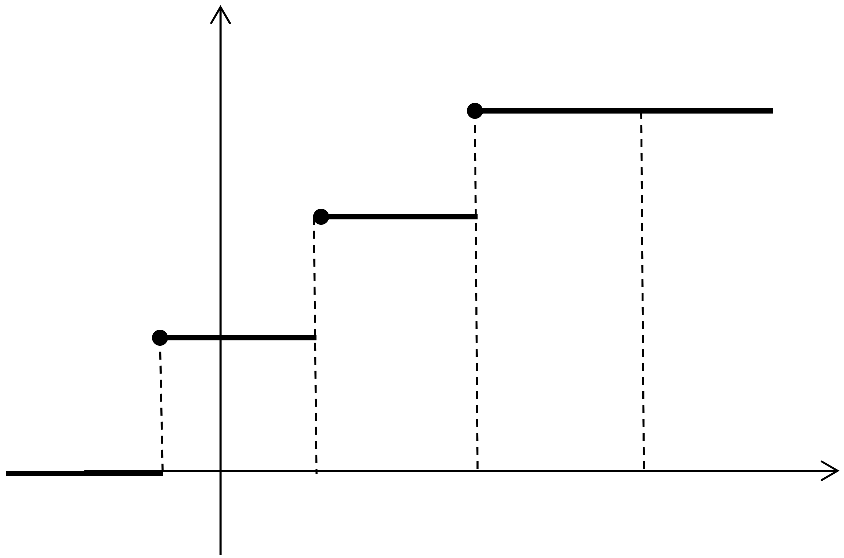
$$\Pr(X < x) = F(x^-).$$

- **Theorem 3.3.4** For every value x,

$$\Pr(X = x) = F(x) - F(x^-)$$

The c.d.f. of a Discrete Distribution

If X has a discrete distribution with the p.f. $f(x)$, then the properties of a c.d.f. imply that $F(x)$ must have the following form: ***F*(*x*) will have a jump by the amount Pr(X=*x*)** **(magnitude** $f(x_i)$ at each possible value x_i of X), and $F(x)$ will be **constant between every pair of successive jumps**.



The c.d.f. of a Continuous Distribution

Theorem 3.3.5 Let X have a continuous distribution, and let $f(x)$ and $F(x)$ denote its p.d.f and c.d.f, respectively. Then ***F* is continuous at every *x*,**

$$F(x) = \int_{-\infty}^{x} f(t)dt, \quad \text{and} \quad \frac{dF(x)}{dx} = f(x),$$

at all x such that f is continuous.

Since the probability of each individual point x is 0, the c.d.f. $F(x)$ will have no jumps.

The c.d.f. of a continuous random variable X can be obtained from the p.d.f. and vice versa.

The c.d.f. of a Continuous Distribution

Ex29 (Book Ex3.3.4) Calculating a p.d.f. from a c.d.f.

The c.d.f. of a *R.V.* is

$$F(x)=\begin{cases}0 & \text{for } x<0,\\ x^{2/3} & \text{for } 0\le x\le 1,\\ 1 & \text{for } x>1.\end{cases}$$

Then its p.d.f. is

$$f(x)=\frac{dF(x)}{dx}=\begin{cases}\frac{2}{3}x^{-1/3} & \text{for } 0<x<1,\\ 0 & \text{otherwise.}\end{cases}$$

The p.d.f. of X can be found at every point other than $x=0$ and $x=1$.

Typical c.d.f. - 1

◆ Ex 30 Binomial c.d.f.

Let X have the Binomial distribution with parameter n and p, Then the c.d.f. of X is

$$F(x)=\begin{cases} 0 & \text{for } x<0, \\ \displaystyle\sum_{i=0}^{floor(x)} \binom{n}{i} p^i(1-p)^{n-i} & \text{for } 0\le x<n, \\ 1 & \text{for } x\ge n. \end{cases}$$

Typical c.d.f. - 2

- **Ex31 Uniform c.d.f.**

Let X have the uniform distribution on the interval [a b], then the c.d.f. of X is

$$F(x)=\begin{cases}0 & \text{for } x<a,\\ \int_a^x \frac{1}{b-a}dy=\frac{x-a}{b-a} & \text{for } a\le x<b,\\ 1 & \text{for } x\ge b.\end{cases}$$

Typical c.d.f. - 3

- **Ex32 Poisson c.d.f.** Let X have the Poisson distribution with parameter λ, Then the c.d.f. of X is

$$F(x)=\begin{cases} 0 & \text{for } x<0, \\ \sum_{i=0}^{floor(x)} \frac{e^{-\lambda}\lambda^i}{i!} & \text{for } x \geq 0. \end{cases}$$

- **Ex 29 Normal c.d.f.**

Let X have the normal distribution with parameter μ and σ, Then the c.d.f. of X is

$$F(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)=\int_{-\infty}^{x}\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(t-\mu)^2}{2\sigma^2}}dt$$

Typical c.d.f. - 4

- **Ex33 Exponential c.d.f.**

Let X have the exponential distribution with parameter β, then the c.d.f. of X is

$$F(x) = \begin{cases} 0 & \text{for } x \le 0, \\ \int_0^x \beta e^{-\beta t} dt = 1 - e^{-\beta x} & \text{for } x > 0. \end{cases}$$

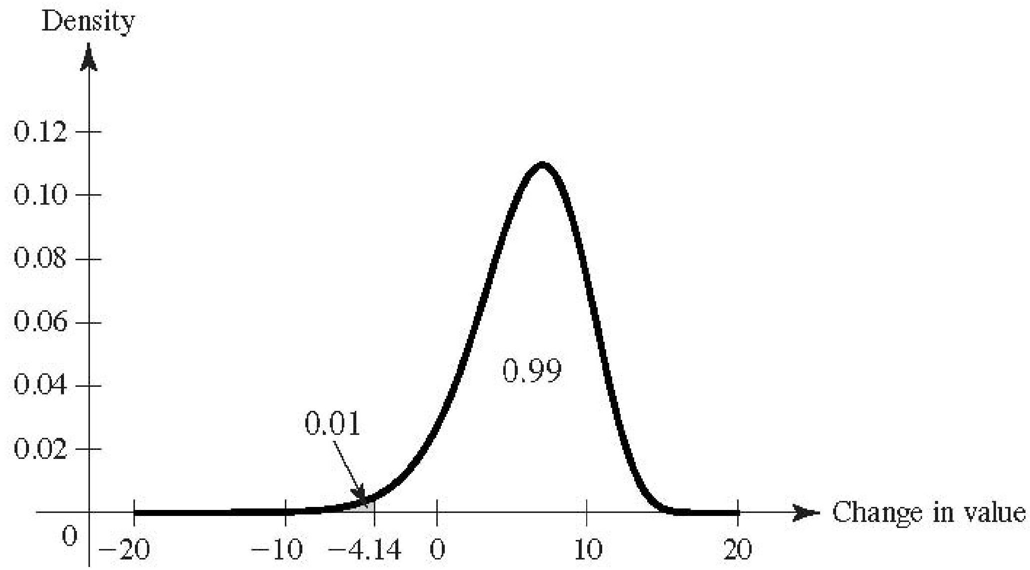
Quantiles/Percentiles - 1

- **Ex34 (Book Ex3.3.5) Fair Bets**. We want to place an even-money bet on X as follows: If $X \leq x_0$, we win one dollar and if $X > x_0$ we lose one dollar. In order to make this bet fair, we need $\Pr(X \leq x_0) = \Pr(X > x_0) = 1/2$.
- If F is a one-to-one function, then F has an inverse function $F^{-1}(x)$,
- $F^{-1}(1/2) = x_0$.
- The value x_0 in this Example is called the 0.5 *quantile* of X or the 50th *percentile* of X because 50% of the distribution of X is at or below x_0.

Quantiles/Percentiles - 2

Definition 3.3.2. Let X be a $R.V.$ with c.d.f. F. For each p strictly between 0 and 1, define $F^{-1}(p)$ to be the smallest value x such that $F(x) \geq p$. Then $F^{-1}(p)$ is called the ***p quantile of X*** or the ***100p percentile of X***. The function F^{-1} defined here on the open interval (0, 1) is called the ***quantile function of X***.

Notice: The above definition extends the concept of inverse function to nondecreasing functions that may be neither continuous nor one-to-one.



Quantiles/Percentiles - 3

- **Ex35 (Book Ex 3.3.8)** Let X have the uniform distribution on the interval [a b], the c.d.f. of X is:

$$F(x) = \begin{cases} 0 & \text{for } x < a, \\ \int_a^x \frac{1}{b-a} dy = \frac{x-a}{b-a} & \text{for } a \le x < b, \\ 1 & \text{for } x \ge b. \end{cases}$$

Q: the quantile function of X?

Sol:

$$\frac{x-a}{b-a} = p \qquad F^{-1}(1/2) = (b+a)/2.$$

$$x = pb + (1-p)a$$

$$F^{-1}(p) = pb + (1-p)a \text{ for } 0<p<1.$$

Quantiles/Percentiles - 4

- **Definition 3.3.3 Median/Quartiles.** The ½ quantile or the 50th percentile of a distribution is called its ***median***. The ¼ quantile or 25th percentile is the ***lower quartile***. The ¾ quantile or 75th percentile is called the ***upper quartile***.

Ex36 (Book 3.3.10) Let X have the uniform distribution on the integers 1,2,3,4, and the c.d.f. of X is

$$F(x) = \begin{cases} 0 & \text{if } x<1, \\ 1/4 & \text{if } 1 \le x < 2, \\ 1/2 & \text{if } 2 \le x < 3, \\ 3/4 & \text{if } 3 \le x < 4, \\ 1 & \text{if } x \ge 4. \end{cases}$$

Q: The 1/2 Quantile is?

2.

Q: What's the median?

Every number in the interval [2,3).

Also, we can say 2.5.

Bivariate Distributions

- **Definition 3.4.1 Joint/Bivariate Distribution.** Let X and Y be $R.V.s$. The ***joint distribution*** or ***bivariate distribution*** of X and Y is the collection of all probabilities of the form $\Pr[(X,Y)\in C]$ for all sets C of pairs of real numbers such that $\{(X,Y)\in C\}$ is an event.
- **Ex37 (Book Ex3.4.1) Demands for Utilities.**

Let X and Y denote the demand for water and electricity, respectively. Define 2 events: $A=\{X\geq 115\}$ and $B=\{Y\geq 110\}$. Define the set of ordered pairs $C=\{(x,y)\colon x\geq 115$ and $y\geq 110\}$ so that the event $\{(X,Y)\in C\}=A\cap B$.

A and B have a joint distribution or bivariate distribution.

Discrete Joint Distributions - 1

- **Definition 3.4.2** Let X and Y be $R.V.$s, and consider the ordered pair (X, Y). If there are only finitely or at most countably many different possible values (x, y) for the pair (X, Y), then we say that X and Y have ***a discrete joint distribution.***
- **Ex38 (Book Ex3.4.2) Theater Patrons.** 10 people is selected at random from a theater with 200 patrons. One $R.V.$ is the number X of people who are over 60 years of age, and another $R.V.$ is the number Y of people who live more than 25 miles from the theater. For each ordered pair (x, y) with $x = 0, \ldots, 10$ and $y = 0, \ldots, 10$. $\Pr\{(X, Y) = (x, y)\}$. X and Y have a discrete joint distribution.

Discrete Joint Distributions - 2

- **Theorem 3.4.1** Suppose that two $R.V.s$ X and Y each have a discrete distribution. Then X and Y have a discrete joint distribution.

- **Definition 3.4.3 Joint Probability Function, p.f.** The joint probability function, or the joint p.f., of X and Y is defined as the function f such that for every point (x, y) in the xy-plane,

$$f(x, y) = \Pr(X = x \text{ and } Y = y).$$

- **Theorem 3.4.2** Let X and Y have a discrete joint distribution. If (x, y) is not one of the possible values of the pair (X,Y), then $f(x, y) = 0$. Also, $\sum_{\text{all } (x,y)} f(x,y) = 1$. For each set C of ordered pairs,

$$\Pr[(X,Y) \in C] = \sum_{(x,y)\in C} f(x,y).$$

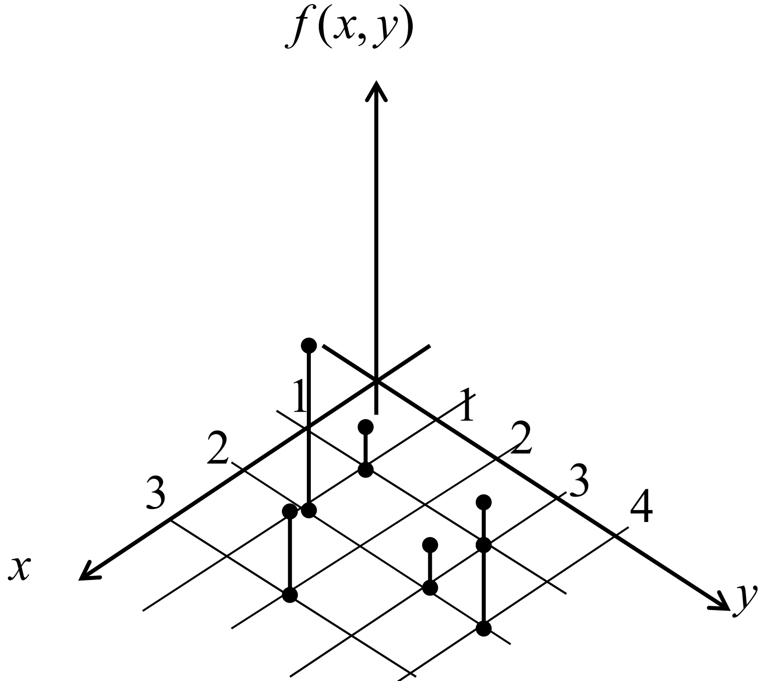
Discrete Joint Distributions - 3

◆ Ex39(Book Ex 3.4.3) Specifying a Discrete Joint Distribution by a Table of Probabilities.

The joint p.f. *f (x, y)*

X	Y			
	1	2	3	4
1	0.1	0	0.1	0
2	0.3	0	0.1	0.2
3	0	0.2	0	0

$$\Pr(X \geq 2 \text{ and } Y \geq 2) = f(2,2) + f(2,3)$$
$$+ f(2,4) + f(3,2) + f(3,3) + f(3,4) = 0.5$$
$$\Pr(X = 1) = \sum_{y=1}^{4} f(1, y) = 0.2$$



The joint p.f. of *X and Y*

Continuous Joint Distributions - 1

- **Definition 3.4.4 Joint p.d.f./Support** 2 $R.V.$s X and Y have a continuous joint distribution if there exists a **nonnegative function *f*** defined over the entire *xy*-plane such that for every subset C of the plane,

$$\Pr[(X,Y)\in C]=\int_C\int f(x,y)dxdy,$$

if the integral exists. The function f is called **the *joint probability density function*** (abbreviated **joint p.d.f.**) of X and Y. The closure of the set $\{(x,y): f(x,y)>0\}$ is called the ***support of (the distribution of) (X, Y)***.

- **Ex40 (Book Ex3.4.5) Demands for Utilities**

The area of S is (150-1)×(200-4)=29204.

$$f(x,y)=\begin{cases}\dfrac{1}{29204} & \text{for } 4\le x\le 200 \text{ and } 1\le y\le 150,\\ 0 & \text{otherwise.}\end{cases}$$

Continuous Joint Distributions - 2

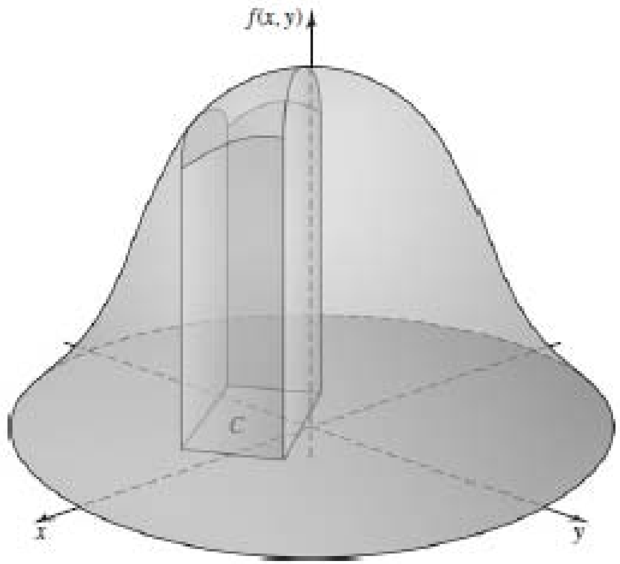
- **Theorem 3.4.3** A joint p.d.f. must satisfy the following two conditions:

$$f(x,y) \geq 0 \quad \text{for } -\infty < x < +\infty \text{ and } -\infty < y < +\infty$$

$$\int_{-\infty}^{+\infty} \int_{-\infty}^{+\infty} f(x,y)\,dx\,dy = 1.$$

Any function that satisfies these two conditions is the joint p.d.f. for some probability distribution.

An example of a joint p.d.f. →



The total volume beneath the surface $f(x,y)$ and above the xy-plane must be 1.

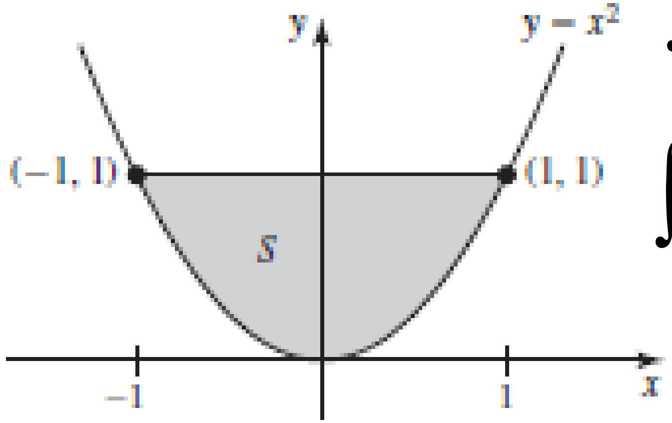
Continuous Joint Distributions - 3

- **Ex41 (Book Ex 3.4.7) Calculating a Normalizing Constant.** Suppose that the joint p.d.f. of X and Y is specified as follows:

$$f(x,y)=\begin{cases} cx^2y & \text{for } x^2 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Determine the value of the constant c.

Sol:



$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dxdy = \int_{-1}^{1}\int_{x^2}^{1} cx^2 y\,dydx$$

or

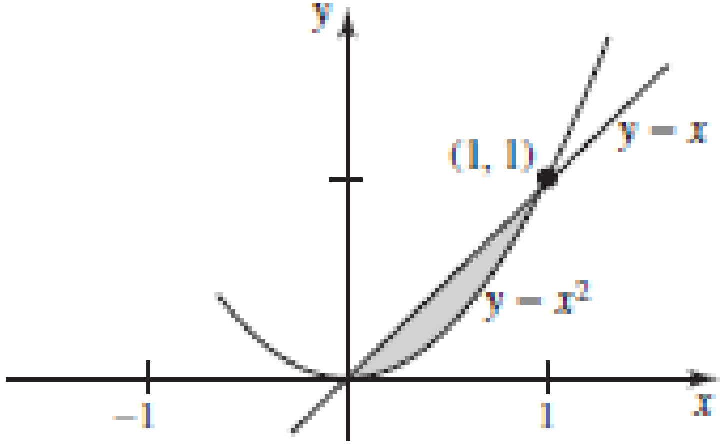
$$\int_{-\infty}^{+\infty}\int_{-\infty}^{+\infty} f(x,y)\,dxdy = \int_{0}^{1}\int_{-\sqrt{y}}^{\sqrt{y}} cx^2 y\,dxdy$$

$$=\frac{4}{21}c = 1 \Rightarrow c = \frac{21}{4}.$$

Continuous Joint Distributions - 4

- **Ex42 (Book Ex 3.4.7) Calculating Probabilities from a Joint p.d.f.** For the joint p.d.f. in Ex41, determine the value $\Pr(X \geq Y)$.

Sol:

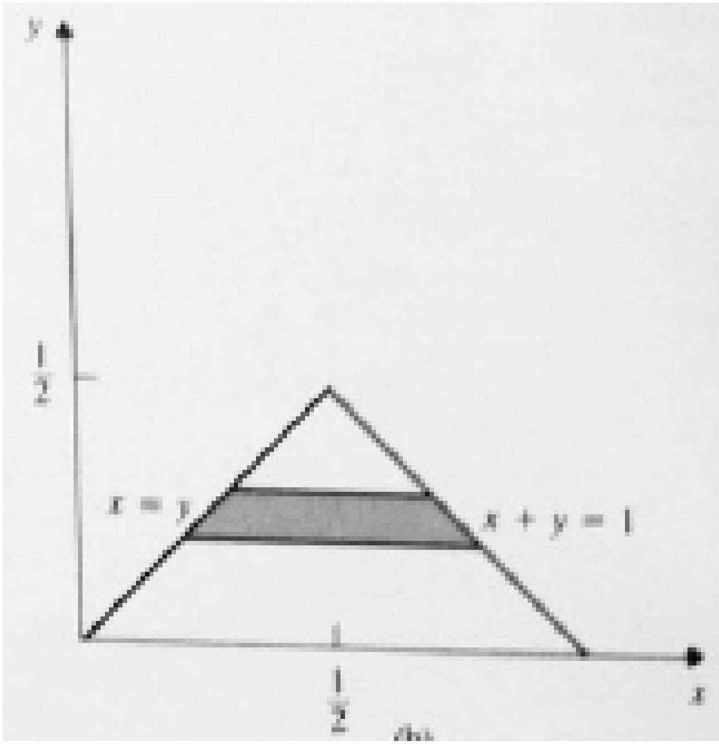


$$\Pr(X \geq Y) = \int_{S_0}\int f(x,y)\,dy\,dx = \int_0^1 \int_{x^2}^{x} \frac{21}{4}x^2 y\,dy\,dx = \frac{3}{20}.$$

or $$\int_{S_0}\int f(x,y)\,dx\,dy = \int_0^1 \int_y^{\sqrt{y}} \frac{21}{4}x^2 y\,dx\,dy = \frac{3}{20}.$$

Continuous Joint Distributions - 5

- Note: be careful with the choice of whether to integrate x or y as the **inner integral**. Also, be careful with the **upper/lower limit** of the definite integral.
- How to choose: **convenience**.



Continuous Joint Distributions - 6

- **Ex43 (Book Ex 3.4.9) Determining a Joint p.d.f. by Geometric Methods.** Suppose that a point (X, Y) is selected at random from inside the circle $x^2+y^2 \le 9$. We shall determine the joint p.d.f. of X and Y.

Sol:

$$\text{because } f(x,y) = \begin{cases} c & \text{for } (x,y) \in S, \\ 0 & \text{otherwise.} \end{cases}$$

$$\text{then } \int_S \int f(x,y)\,dxdy = \quad c \times (\text{area of } S)$$

$$c \times 9\pi = 1 \Rightarrow c = \frac{1}{9\pi}.$$

Mixed Bivariate Distributions - 1

- **Definition 3.4.5 Joint p.f./p.d.f.** Let X and Y be $R.V.s$ such that X is **discrete** and Y is **continuous**. Suppose that there is a function $f(x, y)$ defined on the xy-plane such that, for every pair A and B of subsets of the real numbers,

$$\Pr(X \in A \text{ and } Y \in B) = \int_B \sum_{x \in A} f(x, y)dy,$$

if the integral exists. Then the function f is called the **joint p.f./p.d.f. of X and Y**.

- If X is a discrete $R.V.$ and Y is a continuous $R.V.$, then

$$f(x, y) \geq 0 \text{ for all } x, y,$$

$$\int_{-\infty}^{\infty} \sum_{i=1}^{\infty} f(x_i, y)dy = 1.$$

Mixed Bivariate Distributions - 2

- **Ex44 (Book Ex 3.4.11) A joint p.f./p.d.f.**

Suppose that the joint p.f./p.d.f. of X and Y is

$$f(x,y)=\frac{xy^{x-1}}{3}, \quad \text{for } x=1,2,3 \text{ and } 0<y<1.$$

Check to make sure it satisfies the conditions. Q: If integration is chosen first, over the x values or y values?

$$\sum_{x=1}^{3}\int_0^1 \frac{xy^{x-1}}{3}dy=\sum_{x=1}^{3}\left(\frac{1}{3}y^x\Big|_{y=0}^{y=1}\right)=\sum_{x=1}^{3}\frac{1}{3}=1.$$

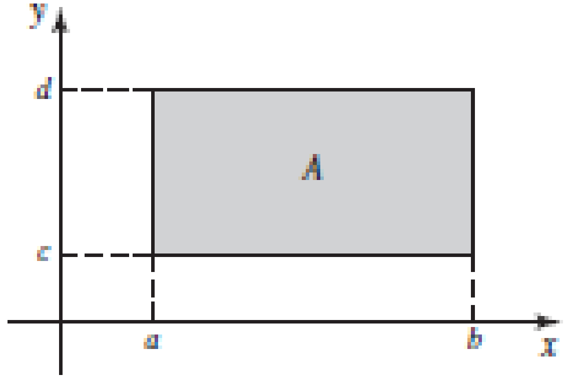
$$\Pr(Y\geq\frac{1}{2} \text{ and } X\geq 2)=\sum_{x=2}^{3}\int_{1/2}^{1}\frac{xy^{x-1}}{3}dy=\sum_{x=2}^{3}\left(\frac{1-(1/2)^x}{3}\right)=0.5417.$$

$$\text{or } \int_{1/2}^{1}[\frac{2}{3}y+y^2]dy=\frac{1}{3}y^2\Big|_{1/2}^{1}+\frac{1}{3}y^3\Big|_{1/2}^{1}= \ 0.5417.$$

Bivariate c.d.f. - 1

◆ **Definition 3.4.6 Joint (Cumulative) Distribution Function/c.d.f.** The joint c.d.f. of two $R.V.s$ X and Y is defined as the function F such that for all values of x and y ($-\infty < x < \infty$ and $-\infty < y < \infty$),

$$F(x, y) = \Pr(X \le x \text{ and } Y \le y)$$



$$\Pr(a < X \le b \text{ and } c < Y \le d)$$

$$= \Pr(a < X \le b \text{ and } Y \le d) - \Pr(a < X \le b \text{ and } Y \le c)$$

$$= [\Pr(X \le b \text{ and } Y \le d) - \Pr(X \le a \text{ and } Y \le d)]$$

$$- [\Pr(X \le b \text{ and } Y \le c) - \Pr(X \le a \text{ and } Y \le c)]$$

$$= F(b,d) - F(a,d) - F(b,c) + F(a,c).$$

Bivariate c.d.f. - 2

- If X and Y have a continuous joint distribution with joint p.d.f. f, then the joint c.d.f. at (x,y) is

$$F(x,y) = \Pr(X \le x \text{ and } Y \le y).$$
$$= \int_{-\infty}^{y} \int_{-\infty}^{x} f(r,s)\,dr\,ds$$

- Given the joint c.d.f., the joint p.d.f. can be derived by using the relations

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \frac{\partial^2 F(x,y)}{\partial y \partial x}$$

at every point (x,y) at which these second-order derivatives exist.

Bivariate c.d.f. - 3

- **Ex45 (Book Ex 3.4.14) Determining a Joint p.d.f. from a Joint c.d.f.** Suppose that X and Y are $R.V.s$ that take values only in the intervals $0 \le X \le 2$ and $0 \le Y \le 2$. Suppose also that the joint c.d.f. of X and Y, for $0 \le x \le 2$ and $0 \le y \le 2$, is as follows:

$$F(x,y) = \frac{1}{16}xy(x+y).$$

determine the joint p.d.f. f of X and Y.

Sol:

$$f(x,y) = \frac{\partial^2 F(x,y)}{\partial x \partial y} = \begin{cases} \frac{1}{8}(x+y) & \text{for } 0 < x < 2 \text{ and } 0 < y < 2, \\ 0 & \text{otherwise.} \end{cases}$$

Bivariate c.d.f. - 4

- **Ex46 (Book Ex 3.4.15) Demands for Utilities.**

$$f(x,y)=\begin{cases}\dfrac{1}{29204} & \text{for } 4\le x\le 200 \text{ and } 1\le y\le 150\\ 0 & \text{otherwise.}\end{cases}$$

Compute the joint c.d.f. for water and electric demand.

$$F(x,y)=\begin{cases}0 & \text{for } x<4 \text{ or } y<1\\ \displaystyle\int_4^x\int_1^y \frac{1}{29204}\,dydx=\frac{(x-4)(y\text{-}1)}{29204} & \text{for } 4\le x\le 200 \text{ and } 1\le y\le 150,\\ \displaystyle\int_4^x\int_1^{150} \frac{1}{29204}\,dydx=\frac{x-4}{196} & \text{for } 4\le x\le 200 \text{ and } y>150,\\ \displaystyle\int_4^{200}\int_1^{y} \frac{1}{29204}\,dydx=\frac{y-1}{149} & \text{for } x>200 \text{ and } 1\le y\le 150,\\ 1 & \text{for } x>200 \text{ and } y>150.\end{cases}$$

Marginal Distributions - 1

- The distribution of one *R.V. X* computed from a joint distribution is called the ***marginal distribution of X***.
- **Theorem 3.4.5** Let X and Y have a joint c.d.f. F. The c.d.f. F_1 of just the single random variable X can be derived from the joint c.d.f. F as $F_1(x) = \lim_{y\to\infty} F(x, y)$. Similarly, the c.d.f. F_2 of Y equals $F_2(y) = \lim_{x\to\infty} F(x, y)$.
- **Definition 3.5.1 Marginal c.d.f./p.f./p.d.f** Suppose that X and Y have a joint distribution. The c.d.f. of X derived by Theorem 3.4.5 is ***the marginal c.d.f. of X***. The p.f. or p.d.f. of X associated with the marginal c.d.f. of X is the ***marginal p.f. or marginal p.d.f. of X***.

Marginal Distributions - 2

- **Ex47 (Book Ex 3.4.14)** Joint c.d.f. is as follows:

$$F(x, y) = \frac{1}{16} xy(x + y).$$

X and Y are $R.V.s$ that take values only in the intervals $0 \le X \le 2$ and $0 \le Y \le 2$. What's the c.d.f. F_1 of X?

Sol: if either $x<0$ or $y<0$, $F(x,y)=0$. If both $x>2$ and $y>2$, $F(x,y)=1$. If $0\le x \le 2$ and $y > 2$, $F(x,y)= F(x,2)$.

By letting $y\to\infty$, the marginal c.d.f. of X is

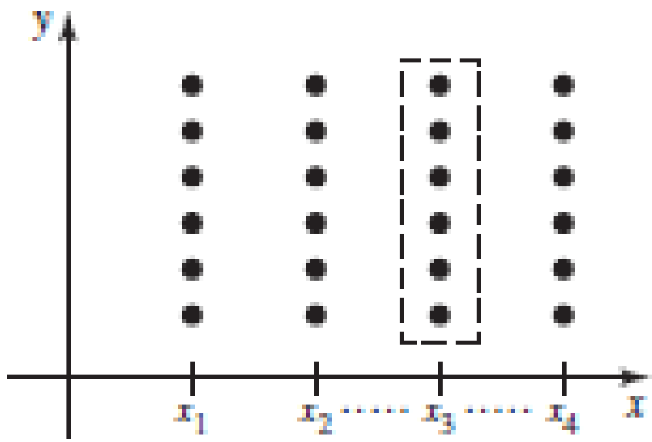
$$F_1(x) = \begin{cases} 0 & \text{for } x<0, \\ \frac{1}{8}x(x+2) & \text{for } 0 \le x \le 2, \\ 1 & \text{for } x>2. \end{cases}$$

Marginal Distributions - 4

◆ **Theorem 3.5.1** If X and Y have a discrete joint distribution for which the joint p.f. is f, then the marginal p.f. f_1 of X is:

$$f_1(x) = \sum_{\text{All } y} f(x, y).$$

Similarly, the marginal p.f. f_2 of Y is $f_2(y) = \sum_{\text{All } x} f(x, y).$



Computing $f_1(x)$ from the joint p.f.

Marginal Distributions - 5

- **Ex48 (Book Ex 3.5.2) Deriving a Marginal p.f. from a Table of Probabilities.**

X	Y				total	$f_1(x)$
	1	2	3	4		
1	0.1	0	0.1	0	0.2	$f_1(1)$
2	0.3	0	0.1	0.2	0.6	$f_1(2)$
3	0	0.2	0	0	0.2	$f_1(3)$
total	0.4	0.2	0.2	0.2	1.0	
$f_2(y)$	$f_2(1)$	$f_2(2)$	$f_2(3)$	$f_2(4)$		

- The name marginal distribution derives from the fact that marginal distributions are the totals that appear in the margins of tables like Table above.

Marginal Distributions - 6

◆ **Theorem 3.5.2** If X and Y have a continuous joint distribution with joint p.d.f. f, then ***the marginal p.d.f. f_1 of X*** is

$$f_1(x) = \int_{-\infty}^{\infty} f(x,y)dy \quad \text{for } -\infty < x < \infty$$

Similarly, ***the marginal p.d.f. f_2 of Y is***

$$f_2(y) = \int_{-\infty}^{\infty} f(x,y)dx \quad \text{for } -\infty < y < \infty$$

Proof: for each x, $\Pr(X \le x)$ can be written by $\Pr[(X,Y)\in C]$, where $C=\{(r,s): r \le x\}$.

$$\Pr[(X,Y)\in C] = \int_{-\infty}^{x}\int_{-\infty}^{\infty} f(r,s)\,ds\,dr$$

$$= \int_{-\infty}^{x}\left[\int_{-\infty}^{\infty} f(r,s)\,ds\right]dr$$

$$= \int_{-\infty}^{x} f_1(r)\,dr = \Pr(X \le x)$$

Marginal Distributions - 3

- **Ex49 (Book Ex 3.5.1) Demands for Utilities.**

$$f(x,y)=\begin{cases}\dfrac{1}{29204} & \text{for } 4\le x\le 200 \text{ and } 1\le y\le 150\\ 0 & \text{otherwise.}\end{cases}$$

It is apparent that the marginal p.d.f. of X is

$$f_1(x)=\begin{cases}\dfrac{1}{196} & \text{for } 4\le x\le 200,\\ 0 & \text{otherwise.}\end{cases}$$

Similarly, the marginal p.d.f. of Y is

$$f_2(y)=\begin{cases}\dfrac{1}{149} & \text{for } 1\le y\le 150,\\ 0 & \text{otherwise.}\end{cases}$$

Marginal Distributions - 7

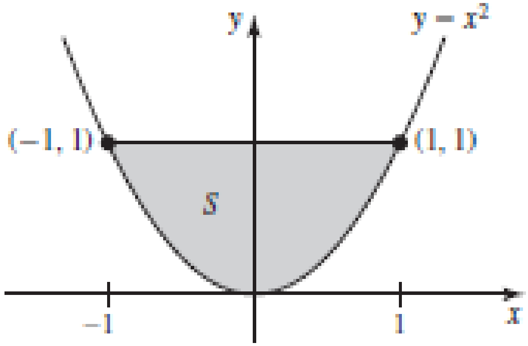
- **Ex50 (Book Ex 3.5.3) Deriving a Marginal p.d.f.**

Suppose that the joint p.d.f. of X and Y is as specified as follows

$$f(x,y)=\begin{cases}\dfrac{21}{4}x^2y & \text{for } x^2\le y\le 1,\\ 0 & \text{otherwise.}\end{cases}$$

Derive the marginal p.d.f. $f_1(x)$ and $f_2(y)$.

Sol:



$$f_1(x)=\int_{-\infty}^{\infty}f(x,y)dy=\int_{x^2}^{1}\frac{21}{4}x^2ydy=\frac{21}{8}x^2(1-x^4).$$

$$f_2(y)=\int_{-\infty}^{\infty}f(x,y)dx=\int_{-\sqrt{y}}^{\sqrt{y}}\frac{21}{4}x^2ydx=\frac{7}{2}y^{5/2}.$$

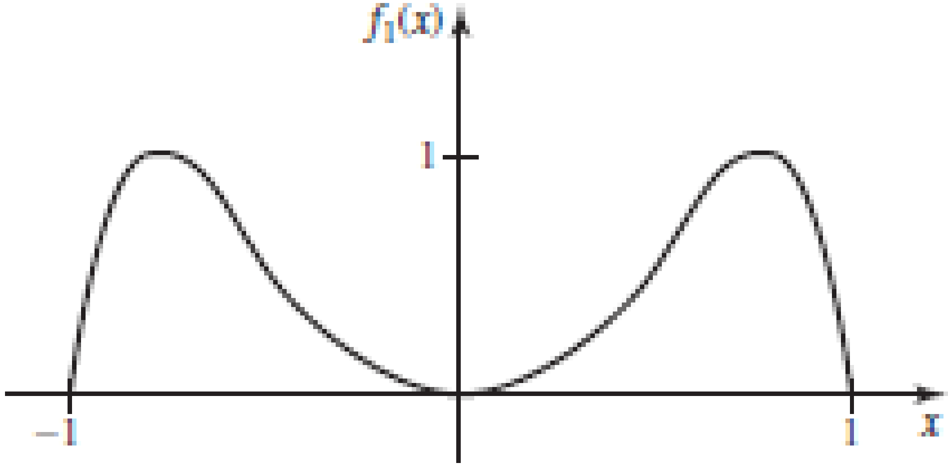
Marginal Distributions - 8

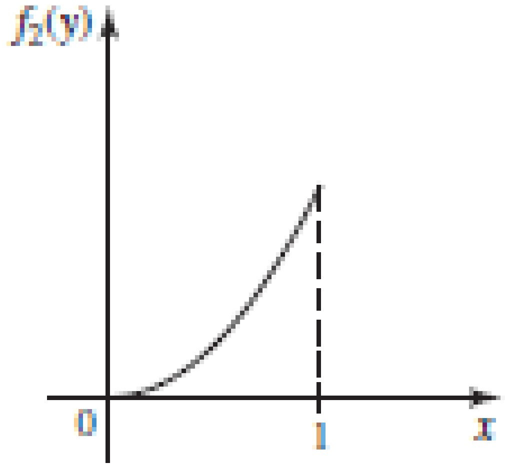
$$f_1(x) = \frac{21}{8} x^2 (1 - x^4).$$

$$\text{for } -1 \le x \le 1.$$

$$f_2(y) = \frac{7}{2} y^{5/2}.$$

$$\text{for } 0 \le y \le 1.$$





Marginal Distributions - 9

- **Theorem 3.5.3** Let f be the joint p.f./p.d.f. of X and Y, with X discrete and Y continuous. Then ***the marginal p.f. of X is:***

$$f_1(x) = \Pr(X = x) = \int_{-\infty}^{\infty} f(x, y)dy \quad \text{for all } x.$$

and ***the marginal p.d.f. of Y is:***

$$f_2(y) = \sum_x f(x, y), \quad \text{for } -\infty < y < \infty.$$

- **Ex51 (Book Ex3.5.4)** The joint p.f./p.d.f. of X and Y

$$f(x, y) = \frac{xy^{x-1}}{3}, \quad \text{for } x = 1, 2, 3 \text{ and } 0 < y < 1.$$

$$f_1(x) = \int_0^1 \frac{xy^{x-1}}{3} dy = \frac{1}{3} \quad \text{for } x = 1, 2, 3.$$

$$f_2(y) = \sum_{x=1}^{3} \frac{xy^{x-1}}{3} = \frac{1}{3} + \frac{2y}{3} + y^2, \quad \text{for } 0 < y < 1.$$

Marginal Distributions - 10

- If X and Y have a continuous joint distribution, then X and Y each have a continuous distribution.
- If X and Y each have a continuous distribution, do X and Y have a continuous joint distribution?

  Probably Not!

  Counter example: $y=f(x)$.

  See Textbook1 P122 Theorem 3.4.4.

X and Y have a discrete joint distribution

$$\Leftrightarrow$$

X and Y each have a discrete distribution

Independent *R.V.s* - 1

◆ **Definition 3.5.2 Independent Random Variables.** Two *R.V.s* X and Y are independent if, for every two sets A and B of real numbers such that $\{X \in A\}$ and $\{Y \in B\}$ are events, $\Pr(X \in A \text{ and } Y \in B) = \Pr(X \in A)\Pr(Y \in B)$.

If X and Y are independent, then **for all real numbers** x and y, it must be true that

$$\Pr(X \le x \text{ and } Y \le y) = \Pr(X \le x)\Pr(Y \le y).$$

◆ **Theorem 3.5.4** Let the joint c.d.f. of X and Y be F, let the marginal c.d.f. of X be F_1 , and let the marginal c.d.f. of Y be F_2 . Then X and Y are independent if and only if, **for all real numbers** x and y,

$$\boxed{F(x, y) = F_1(x)F_2(Y).}$$

Independent *R.V.s* - 2

- **Corollary 3.5.1** Two *R.V.s* X and Y are independent if and only if the following factorization is satisfied for **all real numbers** x and y:

$$\boxed{f(x,y)=f_1(x)f_2(y).}$$

- **Meaning of Independence**

In terms of events, one of two events X occurs does not change the probability that the other one Y occurs. For each y and x such that $\Pr(Y=y)>0$, $\Pr(X=x|Y=y)=\Pr(X=x)$.

- **Ex52 (Book Ex3.5.6)** $F(x,y)=F_1(x)F_2(Y).$

$$F_1(x)=\begin{cases}0 & \text{for } x<4,\\ \dfrac{x}{196} & \text{for } 4\le x\le 200,\\ 1 & \text{for } x>200.\end{cases} F_2(y)=\begin{cases}0 & \text{for } y<1,\\ \dfrac{y}{149} & \text{for } 1\le y\le 150,\\ 1 & \text{for } y>150.\end{cases}$$

Independent *R.V.s* - 3

- **Ex53 (Book Ex3.5.8) Are *X* and *Y* Independent?**

Response (*X*)	Treatment group (*Y*)				total
	Imipramine (1)	Lithium (2)	Combination (3)	Placebo (4)	
Relapse (1)	0.120	0.087	0.146	0.160	0.513
No relapse (2)	0.147	0.166	0.107	0.067	0.487
total	0.267	0.253	0.253	0.227	1.0

- It is seen in the table that $f(1, 2) = 0.087$, while $f_1(1) = 0.513$, and $f_2(2) = 0.253$. Hence, $f(1, 2) \neq f_1(1)f_2(2) = 0.129$. It follows that *X* and *Y* are not independent.

Independent *R.V.s* - 4

- **Ex54 (Book Ex3.5.9) Calculating a Probability.**

Suppose that two measurements X and Y are i.i.d. Suppose that the p.d.f. g of each measurement is as follows:

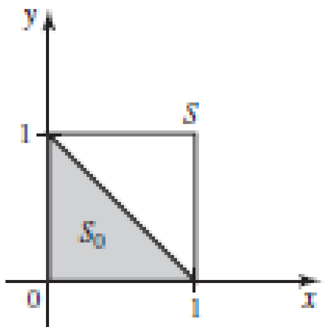
$$g(x) = \begin{cases} 2x & \text{for } 0 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Determine the value of $\Pr(X + Y \le 1)$.

Sol:

$$f(x, y) = g(x)g(y) = \begin{cases} 4xy & \text{for } 0 \le x \le 1 \text{ and } 0 \le y \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Sketch the set S where $f(x, y) > 0$ and the subset S_0 where $x + y \le 1$.



$$\Pr(X + Y \le 1) = \int_{S_0}\int f(x, y)dxdy$$

$$= \int_0^1 \int_0^{1-x} 4xydydx = \frac{1}{6}.$$

Independent *R.V.s* - 5

- **Ex55 (Book Ex3.5.11) Verifying the Factorization of joint p.d.f.** Suppose that the joint p.d.f. of X and Y is

$$f(x,y)=\begin{cases} ke^{-(x+2y)} & \text{for } x\ge 0 \text{ and } y\ge 0,\\ 0 & \text{otherwise.}\end{cases}$$

k is a constant. Determine the marginal p.d.f.'s.

Sol:

$$f_1(x)=\int_0^\infty ke^{-(x+2y)}dy=-\frac{1}{2}ke^{-x}e^{-2y}\Big|_0^\infty=\frac{1}{2}ke^{-x}.$$

$$f_2(y)=\int_0^\infty ke^{-(x+2y)}dx=-ke^{-2y}e^{-x}\Big|_0^\infty=ke^{-2y}.$$

$$\because \int_0^\infty ke^{-2y}dy=1\Rightarrow -\frac{1}{2}ke^{-2y}\Big|_0^\infty=1 \Rightarrow k=2$$

$$\because f(x,y)=f_1(x)f_2(y)\Rightarrow X \text{ and } Y \text{ are independent.}$$

Independent *R.V.s* - 6

- **Note: Separate Functions of Independent *R.V.*s Are Independent.**

If X and Y are independent, then $h(X)$ and $g(Y)$ are independent no matter what the functions h and g are.

Because for every t, the event $\{h(X) \le t\}$ can always be written as $\{X \in A\}$, where $A=\{x: h(x) \le t\}$.

$\{g(Y) \le u\}$ be written as $\{Y \in B\}$, where $B=\{y: g(y) \le u\}$.

As X and Y are independent, we have

$$\Pr(X \in A \text{ and } Y \in B) = \Pr(X \in A)\Pr(Y \in B).$$

$$\Pr[h(X) \le t \text{ and } h(Y) \le u] = \Pr[h(X) \le t]\Pr[h(Y) \le u].$$

Discrete Conditional Distributions-1

- Recall that we've learnt conditional probability

$$P(A\mid B)=\frac{\Pr(A\cap B)}{\Pr(B)}$$

$$P(X=x\mid Y=y)=\frac{\Pr(X=x\text{ and }Y=y)}{\Pr(Y=y)}=\frac{f(x,y)}{f_2(y)}$$

- **Definition 3.6.1 Conditional Distribution/p.f.** Let X and Y have a discrete joint distribution with joint p.f. f. Let f_2 denote the marginal p.f. of Y. For each y such that $f_2(y)>0$, define

$$g_1(x\mid y)=\frac{f(x,y)}{f_2(y)}$$

Then g_1 is the ***conditional p.f. of X given Y***. The discrete distribution whose p.f. is $g_1(.|y)$ is the ***conditional distribution of X given that Y = y.***

Discrete Conditional Distributions-2

- Ex56 (Book Ex3.6.1&3.6.3) Auto Insurance.

Stolen *X*	Brand *Y*					$f_1(x)$ Total
	1	2	3	4	5	
0	0.129	0.298	0.161	0.280	0.108	0.976
1(stolen)	0.010	0.010	0.001	0.002	0.001	0.024
$f_2(y)$ Total	0.139	0.308	0.162	0.282	0.109	1.000

the conditional p.f. of X given Y:

Stolen *X*	Brand *Y*				
	1	2	3	4	5
0	0.928	0.968	0.994	0.993	0.991
1	0.072	0.032	0.006	0.007	0.009

$$g_1(x \mid y) = \frac{f(x,y)}{f_2(y)}$$

Brand 1 is most likely to be stolen.

Discrete Conditional Distributions-3

- **Notice: $g_1(x|y)$ is actually a p.f. as a function of x for each y.**

Let y be such that $f_2(y) > 0$.

$g_1(x \mid y) \geq 0$ for all x and

$$\sum_x g_1(x \mid y) = \frac{\sum_x f(x, y)}{f_2(y)} = \frac{f_2(y)}{f_2(y)} = 1$$

Similarly, we have

$$\sum_y g_2(y \mid x) = 1$$

Stolen *X*	Brand *Y*				
	1	2	3	4	5
0	0.928	0.968	0.994	0.993	0.991
1	0.072	0.032	0.006	0.007	0.009

$g_1(x \mid y)$

Discrete Conditional Distributions-4

- **Ex57 (Book Ex3.6.2) Calculating a Conditional p.f.**

X	Y				total	
	1	2	3	4		$f_1(x)$
1	0.1	0	0.1	0	0.2	$f_1(1)$
2	0.3	0	0.1	0.2	0.6	$f_1(2)$
3	0	0.2	0	0	0.2	$f_1(3)$
total	0.4	0.2	0.2	0.2	1.0	
$f_2(y)$	$f_2(1)$	$f_2(2)$	$f_2(3)$	$f_2(4)$		

Determine the conditional p.f. of Y given that X=2.

Sol: $$g_2(y\,|\,2) = \frac{f(2,y)}{f_1(2)} = \frac{f(2,y)}{0.6}.$$

$g_2(1|2) = 1/2$, $g_2(2|2) = 0$, $g_2(3|2) = 1/6$, $g_2(4|2) = 1/3$.

Continuous Conditional Distribution-1

- **Definition 3.6.2 Conditional p.d.f.** Let X and Y have a continuous joint distribution with joint p.d.f. f and respective marginals f_1 and f_2. Let y be a value such that $f_2(y) > 0$. Then the ***conditional p.d.f. g_1 of X given that $Y = y$*** is defined as follows:

$$g_1(x \mid y) = \frac{f(x, y)}{f_2(y)} \quad \text{for } -\infty < x < \infty$$

For values of y such that $f_2(y) = 0$, we are free to define $g_1(x|y)$ however we wish, so long as $g_1(x|y)$ is a p.d.f. as a function of x.

Continuous Conditional Distribution-2

- **Ex58 (Book Ex3.6.4&Ex3.6.5) Processing Times.**

A manufacturing process consists of two stages. The first stage takes Y minutes, and the whole process takes X minutes (which includes the first Y minutes). Suppose that X and Y have a joint p.d.f. as follows:

$$f(x,y)=\begin{cases} e^{-x} & \text{for } 0\le y\le x<\infty, \\ 0 & \text{otherwise.} \end{cases}$$

Calculate the conditional p.d.f. of X given Y.

Sol: for each y, the possible values of X are all $x\geq y$, so for each $y\geq 0$, $f_2(y)=\int_y^\infty e^{-x}dx=e^{-y}$, and $f_2(y)=0$ for $y<0$.

$$g_1(x|y)=\frac{f(x,y)}{f_2(y)}=\frac{e^{-x}}{e^{-y}}=e^{y-x} \text{ for } x\ge y, \text{ and } g_1(x|y)=0 \text{ for } x<y.$$

$$\Pr(X\ge 9\,|\,Y=4)=\int_9^\infty e^{4-x}dx=e^{-5}=0.0067.$$

Construction of the Joint Distribution-1

- **Theorem 3.6.2 Multiplication Rule for Distributions.**

Let X and Y be $R.V.s$ such that X has p.f. or p.d.f. $f_1(x)$ and Y has p.f. or p.d.f. $f_2(y)$. Also, assume that the conditional p.f. or p.d.f. of X given $Y = y$ is $g_1(x|y)$ while the conditional p.f. or p.d.f. of Y given $X = x$ is $g_2(y|x)$. Then for each y such that $f_2(y) > 0$ and each x,

$$f(x, y) = g_1(x|y)f_2(y),$$

where f is the joint p.f., p.d.f., or p.f./p.d.f. of X and Y. Similarly, for each x such that $f_1(x) > 0$ and each y,

$$f(x, y) = f_1(x)g_2(y|x).$$

Construction of the Joint Distribution-2

- **Ex59 (Book Ex3.6.8) Waiting in a Queue.** Let X be the amount of time that a person has to wait for service in a queue. The faster the server works in the queue, the shorter should be the waiting time. Let Y stand for the rate at which the server works. A common choice of conditional distribution for X given $Y = y$ has conditional p.d.f. for each $y > 0$:

$$g_1(x \mid y) = \begin{cases} ye^{-xy} & \text{for } x \geq 0, \\ 0 & \text{otherwise.} \end{cases}$$

We shall assume that Y has a continuous distribution with p.d.f. $f_2(y) = e^{-y}$ for $y > 0$. Now we can construct the joint p.d.f. of X and Y using Theorem 3.6.2:

$$f(x, y) = g_1(x \mid y) f_2(y) = \begin{cases} ye^{-y(x+1)} & \text{for } x \geq 0, y > 0, \\ 0 & \text{otherwise.} \end{cases}$$

Construction of the Joint Distribution-3

◆ **Theorem 3.6.3 Law of Total Probability for *R.V.s.***

If $f_2(y)$ is the marginal p.f. or p.d.f. of a $R.V.$ Y and $g_1(x|y)$ is the conditional p.f. or p.d.f. of X given $Y = y$, if Y is discrete, then the marginal p.f. or p.d.f. of X is

$$f_1(x) = \sum_y g_1(x \mid y) f_2(y)$$

If Y is continuous, the marginal p.f. or p.d.f. of X is

$$f_1(x) = \int_{-\infty}^{\infty} g_1(x \mid y) f_2(y) dy.$$

Construction of the Joint Distribution-4

- **Theorem 3.6.4 Bayes' Theorem for *R.V.s.***

If $f_2(y)$ is the marginal p.f. or p.d.f. of a *R.V.* Y and $g_1(x|y)$ is the conditional p.f. or p.d.f. of X given $Y = y$, then the conditional p.f. or p.d.f. of Y given X is

$$g_2(y\,|\,x) = \frac{g_1(x\,|\,y)f_2(y)}{f_1(x)} = \frac{g_1(x\,|\,y)f_2(y)}{\sum_y g_1(x\,|\,y)f_2(y)}$$

Similarly, the conditional p.f. or p.d.f. of X given $Y=y$ is

$$g_1(x\,|\,y) = \frac{g_2(y\,|\,x)f_1(x)}{f_2(y)} = \frac{g_2(y\,|\,x)f_1(x)}{\sum_x g_2(y\,|\,x)f_1(x)}$$

Construction of the Joint Distribution-5

- **Ex60 (Book Ex3.6.10)** A point X is chosen from the uniform distribution on the interval(0,1). After the value $X=x$ has been observed ($0<x<1$), a point Y is then chosen from the interval $(x,1)$. What's $g_1(x|y)$?

Sol:

$$f_1(x)=\begin{cases}1 & \text{for } 0<x<1\\ 0 & \text{otherwise.}\end{cases}$$

$$g_2(y\mid x)=\begin{cases}\dfrac{1}{1-x} & \text{for } x<y<1\\ 0 & \text{otherwise.}\end{cases}$$

$$f(x,y)=\begin{cases}\dfrac{1}{1-x} & \text{for } 0<x<y<1,\\ 0 & \text{otherwise.}\end{cases}$$

$$g_1(x\mid y)=\begin{cases}\dfrac{-1}{(1-x)\log(1-y)} & \text{for } 0<x<y,\\ 0 & \text{otherwise.}\end{cases}$$

$$f_2(y)=\int_{-\infty}^{\infty}f(x,y)dx=\int_0^y\frac{1}{1-x}dx=-\log(1-y)\ \text{for } 0<y<1,$$

$$f_2(y)=0\ \text{for } y\le 0 \text{ or } y\ge 1.$$

Construction of the Joint Distribution-6

◆ **Theorem 3.6.5 Independent *R.V.s.***

Suppose that X and Y are two $R.V.s$ having a joint p.f.,p.d.f.,or p.f./p.d.f. f. Then X and Y are independent if and only if for every value of y such that $f_2(y)>0$ and every value of x,

$$g_1(x \mid y) = f_1(x).$$

Proof. Theorem 3.5.4 says that $R.V.s$ X and Y are independent if and only if the following factorization is satisfied for all real numbers x and y:

$$f(x,y) = f_1(x) f_2(y),$$

For $f_2(y)>0$,

$$f_1(x) = \frac{f(x,y)}{f_2(y)} = g_1(x \mid y).$$

Multivariate Distributions - 1

- **Definition 3.7.1 Joint Distribution Function/c.d.f.**

The ***joint c.d.f.*** of n *R.V.s* $X_1, \ldots, X_n$ is the function F whose value at every point $(x_1, \ldots, x_n)$ in n-dimensional space R^n is specified by the relation

$$F(x_1, \cdots, x_n) = \Pr(X_1 \le x_1, X_2 \le x_2, \cdots, X_n \le x_n).$$

- **Ex61 (Book Ex3.7.2) Failure Times.** Suppose that a machine has three parts, and part i will fail at time X_i for $i = 1, 2, 3$. The following function might be the joint c.d.f. of X_1, X_2, and X_3:

$$F(x_1, x_2, x_3) = \begin{cases} (1-e^{-x_1})(1-e^{-2x_2})(1-e^{-3x_3}) & \text{for } x_1, x_2, x_3 \ge 0, \\ 0 & \text{otherwise.} \end{cases}$$

$\underline{X} = (X_1, \ldots, X_n)$ random vector $\underline{X}$ with c.d.f. $F(\underline{x})$.

Multivariate Distributions - 2

- **Definition 3.7.2 Joint Discrete Distribution/p.f.** It is said that n $R.V.s$ $X_1, \ldots, X_n$ have a discrete joint distribution if the random vector $(X_1, \ldots, X_n)$ can have only a finite number or an infinite sequence of different possible values $(x_1, \ldots, x_n)$ in R^n. The joint p.f. of $X_1, \ldots, X_n$ is then defined as the function f such that for every point $(x_1, \ldots, x_n) \in R^n$,

$$f(x_1, \cdots, x_n) = \Pr(X_1 = x_1, \cdots, X_n = x_n).$$

- In vector notation, the joint discrete p.f. becomes

$$f(\underline{x}) = \Pr(\underline{X}).$$

- **Theorem 3.7.1** If $\underline{X}$ has a joint discrete distribution with joint p.f. f, then for every subset $C \subset R^n$

$$\Pr(\underline{X} \in C) = \sum_{\underline{x} \in C} f(\underline{x})$$

Multivariate Distributions - 3

- **Ex62 (Book Ex3.7.1&3.7.3) A Clinical Trial.** m patients are given a treatment, and each patient either recovers or fails to recover. For each $i = 1, \dots, m$, we can let $X_i = 1$ if patient i recovers and $X_i = 0$ if not. There is a $R.V.$ P having a continuous distribution taking values between 0 and 1 such that, if we knew that $P = p$, we would say that the m patients recover with probability p independently of each other.

If $P = p$ is a constant, the joint p.f. of $\underline{X} = (X_1, \dots, X_m)$ is

$$f(\underline{x}) = p^{x_1 + \cdots + x_m}(1-p)^{m - x_1 - \cdots - x_m}$$

for all $x_i \in \{0,1\}$ and 0 otherwise.

Multivariate Distributions - 4

- **Definition 3.7.3 Continuous Distribution/p.d.f.** It is said that n $R.V.s$ $X_1, \ldots, X_n$ have a continuous joint distribution if there is a nonnegative function f defined on R^n such that for every subset $C \subset R^n$

$$\Pr[(X_1, \cdots X_n) \in C] = \int_C \cdots \int f(x_1, \ldots, x_n) dx_1 \cdots dx_n$$

if the integral exits. The function f is called the ***joint p.d.f.*** of $X_1, \ldots X_n$.

In vector notation,

$$\Pr(\underline{X} \in C) = \int_C \cdots \int f(\underline{x}) d\underline{x}$$

Multivariate Distributions - 5

- **Theorem 3.7.2** If the joint distribution of $X_1, \ldots X_n$ is continuous, then the joint p.d.f. f can be derived from the joint c.d.f. F by using the relation

$$f(x_1,\ldots,x_n) = \frac{\partial^n F(x_1,\ldots,x_n)}{\partial x_1 \cdots \partial x_n}$$

at all points $(x_1, \ldots, x_n)$ at which the derivative in this relation exists.

- **Ex63 (Book Ex3.7.2&3.7.4) Failure Times.** Find the joint p.d.f. for the three $R.V.$s in Ex 3.7.2.

$$f(x_1,x_2,x_3) = \begin{cases} 6e^{-x_1-2x_2-3x_3} & \text{for } x_1,x_2,x_3 > 0, \\ 0 & \text{otherwise.} \end{cases}$$

- Note: even if each of $X_1, \ldots X_n$ has a continuous distribution, the vector $\underline{X}$ might not have a continuous joint distribution. Check Theorem3.4.4.

Mixed Distributions

- **Definition 3.7.4 Joint p.f./p.d.f.** Let $X_1, \ldots X_n$ be $R.V.s$, Some of which have a continuous joint distribution and some of which have discrete distributions; their joint distribution would then be represented by a function f that we call the ***joint p.f./p.d.f.***

- **Ex 64 (Book Ex3.7.6) Arrivals at a Queue.** Let Z stand for the rate at which customers are served. Let Y stand for the rate at which customers arrive at the queue. Finally, let W stand for the number of customers that arrive during one day. Then W is discrete while Y and Z could be continuous $R.V.s$. A possible joint p.f./p.d.f. is

$$f(y,z,w) = \begin{cases} 6e^{-3z-10y}(8y)^w / w! & \text{for } z, y > 0 \text{ and } w\text{=0,1,...,} \\ 0 & \text{otherwise.} \end{cases}$$

Marginal Distributions - 1

- **Deriving a Marginal p.d.f.** If the joint distribution of n random variables $X_1, \ldots, X_n$ is known, then the marginal distribution of each single random variable X_i can be derived from this joint distribution.

$$f_1(x_1) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n-1} f(x_1, \cdots, x_n) dx_2 \cdots dx_n.$$

- More generally, the marginal joint p.d.f. of any k of the n random variables $X_1, \ldots, X_n$ can be found by integrating the joint p.d.f. over all possible values of the other $n-k$ variables.

$$f_{24}(x_2, x_4) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f(x_1, x_2, x_3, x_4) dx_1 dx_3$$

Marginal Distributions - 2

- **Deriving a Marginal c.d.f.** Consider now a joint distribution for which the joint c.d.f. of $X_1, \ldots, X_n$ is F. The marginal c.d.f. F_1 of X_1 can be obtained from the following relation:

$$F_1(x_1) = \Pr(X_1 \le x_1) = \Pr(X_1 \le x_1, X_2 < \infty, \ldots, X_n < \infty)$$
$$= \lim_{x_2, \ldots, x_n \to \infty} F(x_1, x_2, \ldots, x_n).$$

- **Ex 65 (Book Ex3.7.10 & Ex3.7.11) Failure Times.**

Find the marginal c.d.f. of X_1.
Let x_2 and x_3 go to ∞. $F_1(x_1) = 1 - e^{-x_1}$ for $x_1 \ge 0$ and 0 otherwise.
How about the marginal bivariate c.d.f. of X_1&X_3?

$$F(x_1, x_3) = \begin{cases} (1-e^{-x_1})(1-e^{-3x_3}) & \text{for } x_1, x_3 \ge 0, \\ 0 & \text{otherwise.} \end{cases}$$

Independent $R.V.s$

- **Definition 3.7.5 Independent $R.V.s$.** It is said that n $R.V.s$ $X_1, \ldots, X_n$ are ***independent*** if, for every n sets $A_1, \ldots, A_n$ of real numbers,
$$\Pr(X_1 \in A_1, X_2 \in A_2, \cdots, X_n \in A_n) = \Pr(X_1 \in A_1)\Pr(X_2 \in A_2)\cdots\Pr(X_n \in A_n).$$

- **Theorem 3.7.3** The variables $X_1, \ldots, X_n$ are independent if and only if, for all points $(x_1, x_2, \cdots, x_n) \in R^n$,
$$F(x_1, x_2, \cdots, x_n) = F_1(x_1)F_2(x_2)\cdots F_n(x_n).$$

- **Theorem 3.7.4** The variables $X_1, \ldots, X_n$ are independent if and only if, for all points $(x_1, x_2, \cdots, x_n) \in R^n$,
$$f(x_1, x_2, \cdots, x_n) = f_1(x_1)f_2(x_2)\cdots f_n(x_n).$$

Conditional Distributions - 1

- Suppose that n $R.V.s$ $X_1, \ldots, X_n$ have a continuous joint distribution for which the joint p.d.f. is f and that f_0 denotes the marginal joint p.d.f. of the $k<n$ $R.V.s$ $X_1, \ldots, X_k$. Then $x_1, \ldots, x_k$ such that $f_0(x_1, \ldots, x_k) > 0$, the conditional p.d.f. of $(X_{k+1}, \ldots, X_n)$ given that $X_1 = x_1, \ldots, X_k = x_k$ is defined as follows:

$$g_{k+1,\cdots n}(x_{k+1}, \cdots, x_n \mid x_1, \cdots, x_k) = \frac{f(x_1, x_2, \cdots, x_n)}{f_0(x_1, \cdots, x_k)}$$

Conditional Distributions - 2

- **Definition 3.7.7 Conditional p.f., p.d.f., or p.f./p.d.f.**

Suppose that the $R.V.$ $\underline{X} = (X_1, \ldots, X_n)$ is divided into two subvectors $\underline{Y}$ and $\underline{Z}$, where $\underline{Y}$ is a k-dimensional random vector comprising k of the n $R.V.$s in $\underline{X}$, and $\underline{Z}$ is an $(n-k)$-dimensional random vector comprising the other $n-k$ $R.V.$s in $\underline{X}$. Suppose also that the n-dimensional joint p.f., p.d.f., or p.f./p.d.f. of $(\underline{Y}, \underline{Z})$ is f and that the marginal $(n-k)$-dimensional p.f., p.d.f., or p.f./p.d.f. of $\underline{Z}$ is f_2. Then for every given point $\underline{z} \in R^{n-k}$ such that $f_2(\underline{z}) > 0$, the conditional k-dimensional p.f., p.d.f., or p.f./p.d.f. g_1 of Y given $\underline{Z} = \underline{z}$ is defined as follows:

$$g_1(\underline{y} \mid \underline{z}) = \frac{f(\underline{y}, \underline{z})}{f_2(\underline{z})} \quad \text{for } \underline{y} \in R^k. \text{ or } f(\underline{y}, \underline{z}) = g_1(\underline{y} \mid \underline{z}) f_2(\underline{z}).$$

Conditional Distributions - 3

- **Definition 3.7.8 Conditionally Independent *R.V.*s.**

Let $\underline{Z}$ be a random vector with joint p.f., p.d.f., or p.f./p.d.f. $f_0(\underline{z})$. Several *R.V.*s $X_1, \ldots, X_n$ are conditionally independent given $\underline{Z}$ if, for all $\underline{z}$ such that $f_0(\underline{z}) > 0$, we have

$$g(\underline{x} \mid \underline{z}) = \prod_{i=1}^{n} g_i(x_i \mid \underline{z}).$$

where $g(\underline{x}|\underline{z})$ stands for the conditional multivariate p.f., p.d.f., or p.f./p.d.f. of $\underline{X}$ given $\underline{Z} = \underline{z}$ and $g_i(x_i|\underline{z})$ stands for the conditional univariate p.f. or p.d.f. of X_i given $\underline{Z} = \underline{z}$.

Please self-study Histograms.

Ex 66 (Book Ex3.7.15 & Ex3.7.16) Suppose that Z is a *R.V.* for which the p.d.f. f_0 is as follows:

$$f_0(z)=\begin{cases}2e^{-2z} & \text{for } z>0,\\ 0 & \text{otherwise.}\end{cases} \qquad g(x\mid z)=\begin{cases}ze^{-zx} & \text{for } x>0,\\ 0 & \text{otherwise.}\end{cases}$$

for every given value $Z=z>0$ two other *R.V.*s X_1 and X_2 are i.i.d. and the conditional p.d.f. of each of them is ↑, determine the marginal joint p.d.f. of (X_1,X_2).

$$g_{12}(x_1,x_2\mid z)=\begin{cases}z^2e^{-z(x_1+x_2)} & \text{for } x_1,x_2>0,\\ 0 & \text{otherwise.}\end{cases}$$

$$f(z,x_1,x_2)=f_0(z)g_{12}(x_1,x_2\mid z)=2z^2e^{-z(2+x_1+x_2)} \text{ for } x_1,x_2,z>0.$$

the marginal joint p.d.f. of (X_1,X_2) is:

$$\boxed{f_{12}(x_1,x_2)=\int_0^\infty f(z,x_1,x_2)dz=4/(2+x_1+x_2)^3 \text{ for } x_1,x_2>0.}$$

$$\Pr(X_1+X_2<4)=\int_0^4\int_0^{4-x_2}\frac{4}{(2+x_1+x_2)^3}dx_1dx_2=\frac{4}{9}$$

For every value of z, the conditional p.d.f. of Z given $X_1=x_1, X_2=x_2$ is:

$$g_0(z \mid x_1, x_2)$$

$$= \frac{f(z, x_1, x_2)}{f_{12}(x_1, x_2)} = \begin{cases} \frac{1}{2}(2 + x_1 + x_2)^3 z^2 e^{-z(2+x_1+x_2)} & \text{for } z>0, \\ 0 & \text{otherwise.} \end{cases}$$

$$\Pr(Z \le 1 \mid X_1 = 1, X_2 = 4)$$

$$= \int_0^1 g_0(z \mid 1, 4) dz$$

$$= \int_0^1 171.5 z^2 e^{-7z} dz = 0.9704.$$

Conditional Distributions - 4

◆ **Theorem 3.7.5 Law of Total Probability and Bayes' Theorem** Assume that $\underline{Y}$ is a k-dimensional random vector, and $\underline{Z}$ is an $(n-k)$-dimensional random vector, and the conditional p.f., p.d.f, or p.f./p.d.f. of $\underline{Y}$ given $\underline{Z}=\underline{z}$ is $g_1(\underline{y}|\underline{z})$.

If $\underline{Z}$ has a continuous joint distribution, the marginal p.d.f. of $\underline{Y}$ is

$$f_1(\underline{y}) = \underbrace{\int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty}}_{n-k} g_1(\underline{y} \mid \underline{z}) f_2(\underline{z}) d\underline{z},$$

and the conditional p.d.f. of $\underline{Z}$ given $\underline{Y}=\underline{y}$ is

$$g_2(\underline{z} \mid \underline{y}) = \frac{g_1(\underline{y} \mid \underline{z}) f_2(\underline{z})}{f_1(\underline{y})}.$$

Functions of a *R.V.*

- **Ex67 (Book Ex3.8.1) Distance from the Middle.**

Let X have the uniform distribution on the integers 1, 2, . . . , 9. Suppose that we are interested in how far X is from the middle of the distribution, namely, 5.

We could define $Y = |X - 5|$ and compute probabilities such as

$$\Pr(Y = 1) = \Pr(X \in \{4, 6\}) = \frac{2}{9}.$$

- **Theorem 3.8.1 Function of a Discrete *R.V.*** Let X have a discrete distribution with p.f. f, and let $Y = r(X)$ for some function of r defined on the set of possible values of X. For each possible value y of Y, the p.f. g of Y is

$$g(y) = \Pr(Y = y) = \Pr[r(X) = y] = \sum_{x:\ r(x)=y} f(x)$$

R.V. with a Discrete Distribution

- **Ex68** Suppose that a *R.V.* X can have each of the seven values −3, −2, −1, 0, 1, 2, 3 with equal probability. Determine the p.f. of $Y = X^2 - X$.

Sol: we have that

Since the probability of each value of X is 1/7, the probability distribution of Y is

x	y
-3	12
-2	6
-1	2
0	0
1	0
2	2
3	6

The p.f. of Y is

$$f(y)=\begin{cases}\frac{2}{7}, & \text{for } y=0\\ \frac{2}{7}, & \text{for } y=2\\ \frac{2}{7}, & \text{for } y=6\\ \frac{1}{7}, & \text{for } y=12\\ 0 & \text{otherwise}\end{cases}$$

y	$g(y)$
0	$\frac{2}{7}$
2	$\frac{2}{7}$
6	$\frac{2}{7}$
12	$\frac{1}{7}$

$R.V.$ with a Continuous Distribution - 1

- **Ex69 (Book Ex3.8.3) Average Waiting Time.** Let Z be the rate at which customers are served in a queue, and suppose that Z has a continuous c.d.f. F. The average waiting time is $Y = 1/Z$. If we want to find the c.d.f. G of Y, we can write

$$G(y) = \Pr(Y \le y) = \Pr(\frac{1}{Z} \le y) = \Pr(Z \ge \frac{1}{y}) = \Pr(Z > \frac{1}{y}) = 1 - F(\frac{1}{y}).$$

Suppose that the p.d.f. of X is f and that another $R.V.$ is defined as $Y=r(X)$. For each real number y, the c.d.f. $G(y)$ of Y can be derived as follows:

$$G(y) = \Pr(Y \le y) = \Pr[r(x) \le y] = \int_{\{x:\, r(x) \le y\}} f(x)dx$$

If Y is continuous, its p.d.f.

$d(y) = dG(y)/dy$ at every point y at which G is differentiable.

R.V. with a Continuous Distribution - 2

- **Ex70 (Book Ex3.8.4) Deriving the p.d.f. of X^2 when X Has a Uniform Distribution.** Suppose that X has the uniform distribution on the interval $[-1, 1]$, so

$$f(x) = \begin{cases} 1/2 & \text{for } -1 \le x \le 1, \\ 0 & \text{otherwise.} \end{cases}$$

Determine the p.d.f. of the random variable $Y = X^2$.

Sol: since $Y = X^2$, $0 \le Y \le 1$. Thus, for each value of Y such that $0 \le y \le 1$, the c.d.f. $G(y)$ of Y is

$$G(y) = \Pr(Y \le y) = \Pr(X^2 \le y) = \Pr(-\sqrt{y} \le X \le \sqrt{y}) = \int_{-\sqrt{y}}^{\sqrt{y}} f(x)dx = \sqrt{y}.$$

For $0<y<1$, it follows that the p.d.f. $g(y)$ of Y is

$$g(y) = \frac{dG(y)}{dy} = \frac{1}{2y^{1/2}}.$$

The p.d.f. of Y is unbounded in the neighborhood of $y=0$.

R.V. with a Continuous Distribution - 3

- **Theorem 3.8.2 Linear Function.** Suppose that X is a *R.V.* for which the p.d.f. is f and that $Y = aX + b$ ($a \neq 0$). Then the p.d.f. of Y is

$$g(y) = \frac{1}{|a|} f\left(\frac{y-b}{a}\right) \quad \text{for} \quad -\infty < y < \infty.$$

Proof . If $a>0$

$$G(y) = \Pr(Y \le y) = \Pr(aX + b \le y) = \Pr\left(X \le \frac{y-b}{a}\right) = F\left(\frac{y-b}{a}\right).$$

Obtain the p.d.f. of Y by differentiating with respect to y.

$$\frac{dG(y)}{dy} = \frac{dF\left(\frac{y-b}{a}\right)}{dy} = \frac{1}{a} f\left(\frac{y-b}{a}\right).$$

If $a<0$

$$G(y)=\Pr(Y\le y)=\Pr(aX+b\le y)=\Pr(X\ge\frac{y-b}{a})=1-F(\frac{y-b}{a}).$$

Obtain the p.d.f. of Y by differentiating with respect to y.

$$\frac{dG(y)}{dy}=\frac{d[1-F(\frac{y-b}{a})]}{dy}=\frac{1}{-a}f(\frac{y-b}{a}).$$

Thus the p.d.f. of Y is

$$g(y)=\frac{1}{|a|}f(\frac{y-b}{a}) \quad \text{for} \quad -\infty<y<\infty.$$

R.V. with a Continuous Distribution - 4

- More generally, if the equation $r(x)=y$ has n solutions,

$$g(y)=\sum_{k=1}^{n}\frac{f(x)}{|dy/dx|}\Bigg|_{x=x_k}$$

- **Ex71** Suppose X is a standard normal *R.V.* Let $Y=X^2$.

Find the p.d.f. of Y.

Sol:

$$f(x)=\frac{1}{\sqrt{2\pi}}\exp(-\frac{1}{2}x^2)\text{ for }-\infty<x<\infty.$$

The equation that $y=x^2$ has 2 solutions: $x_1=\sqrt{y}$ and $x_2=-\sqrt{y}$.

So $g(y)$ has two terms for $y\geq 0$,

$$g(y)=\frac{f(\sqrt{y})}{2\sqrt{y}}+\frac{f(-\sqrt{y})}{2\sqrt{y}}=\frac{1}{\sqrt{2\pi y}}\exp(-\frac{y}{2}).$$

Otherwise $g(y)=0$.

The probability Integral Transformation

- **Theorem 3.8.3** Let X have a continuous c.d.f. F, and let $Y=F(X)$. This transformation from X to Y is called ***the probability integral transformation***. The distribution of Y is the uniform distribution on the interval [0,1].

- **Ex72 (Book Ex3.8.5)** Let X be a continuous $R.V.$ with p.d.f. $f(x) = \exp(-x)$ for $x>0$ and 0 otherwise. The c.d.f. of X is $F(x) = 1-\exp(-x)$ for $x>0$ and 0 otherwise. we will find the distribution of $Y = F(X)$.

Sol: for $0<y<1$， the c.d.f. of Y is

$$G(y) = \Pr(Y \le y) = \Pr(1-\exp(-X) \le y) = \Pr[X \le -\ln(1-y)]$$

$$= F[-\ln(1-y)] = 1-\exp\{-[-\ln(1-y)]\} = y.$$

which is the c.d.f. of the uniform distribution on the Interval[0,1].

Functions of two or more *R.V.*s - 1

- **Ex73 (Book Ex3.9.1&3.9.2) Bull Market.** 3 firms, each has 10 funds. Let $X_i=1$ if fund i performs better than the standard and $X_i=0$ otherwise. We are interested in

$$Y_1=X_1+\ldots+X_{10},$$
$$Y_2=X_{11}+\ldots+X_{20},$$
$$Y_3=X_{21}+\ldots+X_{30}.$$

What is the joint p.f. g of (Y_1,Y_2,Y_3) at the point (3,5,8) if all possible values of $(X_1,\ldots, X_{30})$ are equally likely.

Sol: we can define the set A as

$$A=\{(x_1,\ldots,x_{30}):x_1+\ldots+x_{10}=3,x_{11}+\ldots+x_{20}=5,x_{21}+\ldots+x_{30}=8\}.$$

How many points in A?

$$\binom{10}{3}\binom{10}{5}\binom{10}{8}=1{,}360{,}800$$

$$g(3, 5, 8) = 1{,}360{,}800/\ 2^{30} = 1.27 \times 10^{-3}.$$

Functions of two or more *R.V.*s - 2

- **Theorem 3.9.1 Functions of Discrete *R.V.*s.**

Suppose that n *R.V.*s $X_1, \ldots, X_n$ have a discrete joint distribution for which the joint p.f. is f, and that m functions $Y_1, \ldots, Y_m$ of these n *R.V.*s are as follows:

$$\begin{aligned} Y_1 &= r_1(X_1, \ldots, X_n), \\ Y_2 &= r_2(X_1, \ldots, X_n), \\ &\cdots \\ Y_m &= r_m(X_1, \ldots, X_n). \end{aligned} \qquad \Bigg| \qquad \begin{aligned} r_1(x_1, \ldots, x_n) &= y_1, \\ r_2(x_1, \ldots, x_n) &= y_2, \\ &\cdots \\ r_m(x_1, \ldots, x_n) &= y_m. \end{aligned}$$

For given values $y_1, \ldots, y_m$ of the m *R.V.*s $Y_1, \ldots, Y_m$, let A denote the set of all points $(x_1, \ldots, x_n)$ such that ↑

Then the value of the joint p.f. g of $Y_1, \ldots, Y_m$ is specified at the point $(y_1, \ldots, y_m)$ by the relation

$$g(y_1, \ldots, y_m) = \sum_{(x_1, \ldots x_n) \in A} f(x_1, \ldots x_n)$$

Functions of two or more *R.V.*s - 3

◆ **Theorem 3.9.3 Brute-Force Distribution of a Function.** Suppose that the joint p.d.f. of $\underline{X} = (X_1, \ldots, X_n)$ is $f(\underline{x})$ and that $Y = r(\underline{X})$. For each real number y, define $A_y = \{\underline{x} : r(\underline{X}) \le y\}$. Then the c.d.f. $G(y)$ of Y is

$$G(y) = \int_{A_y} \cdots \int f(\underline{x}) d\underline{x}.$$

Proof: $G(y) = \Pr(Y \le y) = \Pr[r(\underline{X}) \le y] = \Pr(\underline{X} \in A_y).$

If the distribution of Y is continuous, then the p.d.f. of Y can be found by differentiating the c.d.f. G(y).

Functions of two or more *R.V.*s - 4

- **Ex74 (Book Ex3.9.4) Total Service Time.**

Suppose that the first two customers in a queue plan to leave together. Let X_i be the time it takes to serve customer i for $i = 1, 2$. Suppose also that X_1 and X_2 are independent *R.V.*s with common distribution having p.d.f. $f(x) = 2e^{-2x}$ for $x > 0$ and 0 otherwise. Since the customers will leave together, they are interested in the total time it takes to serve both of them, namely, $Y = X_1 + X_2$.

Find the c.d.f of Y and the p.d.f *of* Y.

Sol: let $A_y = \{(x_1, x_2) : x_1 + x_2 \le y\}.$

Then $Y \le y$ if and only if $(X_1, X_2) \in A_y$.

For $y>0$, $G(y)$=?

$$G(y) = \Pr((X_1, X_2) \in A_y) = \int_0^y \int_0^{y-x_2} 4e^{-2x_1 - 2x_2} dx_1 dx_2$$

$$= 1 - e^{-2y} - 2ye^{-2y}.$$

Taking the derivative of $G(y)$ with respect to y, we get the p.d.f.

$$g(y) = \begin{cases} \frac{d}{dy}[1 - e^{-2y} - 2ye^{-2y}] = 4ye^{-2y} & \text{for } y > 0 \\ 0 & \text{otherwise.} \end{cases}$$

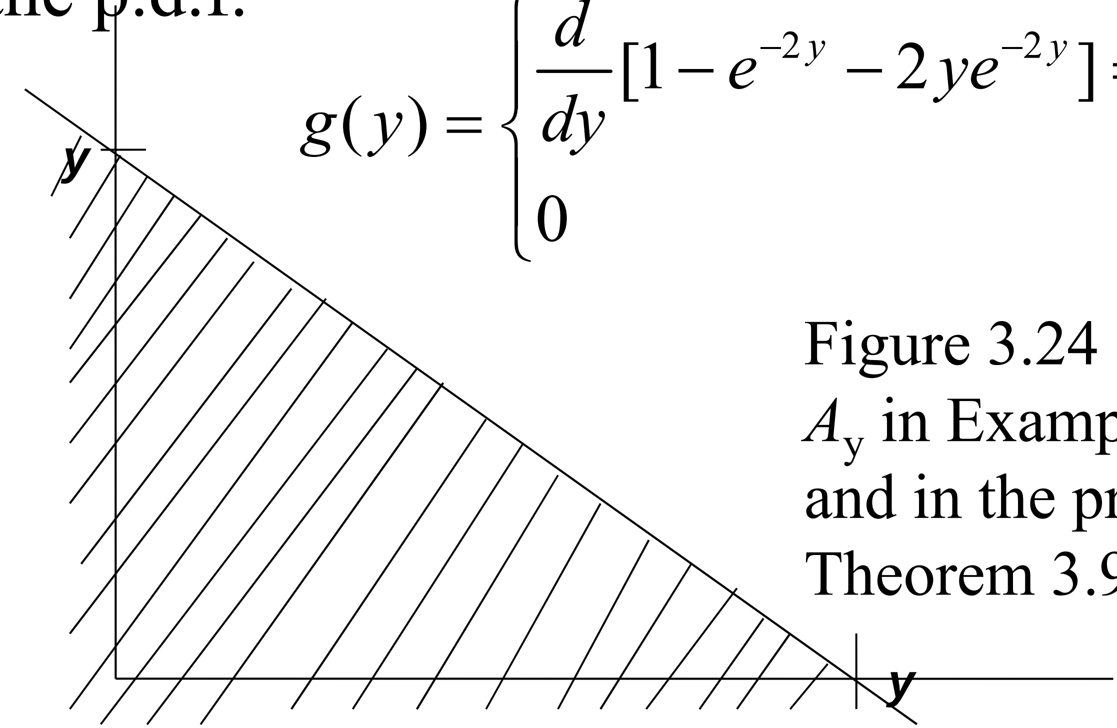


Figure 3.24 The set A_y in Example 3.9.4 and in the proof of Theorem 3.9.4.

Functions of two or more *R.V.*s - 5

Theorem 3.9.4 Linear Function of Two *R.V.*s. Let X_1 and X_2 have joint p.d.f. $f(x_1, x_2)$, and let $Y = a_1X_1 + a_2X_2 + b$ with $a_1 \neq 0$. Then Y has a continuous distribution whose p.d.f. is

$$g(y) = \int_{-\infty}^{\infty} f\left(\frac{y-b-a_2x_2}{a_1}, x_2\right)\frac{1}{|a_1|}dx_2$$

Proof: for each y, let $A_y = \{(x_1,x_2) : a_1x_1 + a_2x_2 + b \le y\}$.
Assume $a_1>0$, the other case is similar.

$$G(y) = \int_{A_y}\int f(x_1,x_2)dx_1dx_2 = \int_{-\infty}^{\infty}\int_{-\infty}^{(y-b-a_2x_2)/a_1} f(x_1,x_2)dx_1dx_2$$

Make $z = a_1x_1 + a_2x_2 + b, dx_1 = dz / a_1$.

$$G(y) = \int_{-\infty}^{\infty}\int_{-\infty}^{y} f(\frac{z-b-a_2x_2}{a_1}, x_2)\frac{1}{a_1}dzdx_2 = \int_{-\infty}^{y}\int_{-\infty}^{\infty} f(\frac{z-b-a_2x_2}{a_1}, x_2)\frac{1}{a_1}dx_2dz$$

$$\because G(y) = \int_{-\infty}^{y} g(z)dz. \quad \therefore g(y) = \int_{-\infty}^{\infty} f\left(\frac{y-b-a_2x_2}{a_1}, x_2\right)\frac{1}{a_1}dx_2$$

Functions of two or more *R.V.*s - 6

◆ **Definition 3.9.1 Convolution.** Let X_1 and X_2 be independent continuous *R.V.*s and let $Y = X_1 + X_2$. The distribution of Y is called the ***convolution*** of the distributions of X_1 and X_2. The p.d.f. of Y is sometimes called the ***convolution*** of the p.d.f.'s of X_1 and X_2. Theorem 3.9.4 says that the p.d.f. of $Y = X_1 + X_2$ is

$$g(y) = \int_{-\infty}^{\infty} f_1(y-z) f_2(z)\,dz.$$

$$\text{or}\quad g(y) = \int_{-\infty}^{\infty} f_1(z) f_2(y-z)\,dz.$$

$$g(y) = \int_{-\infty}^{\infty} f\left(\frac{y-b-a_2x_2}{a_1}, x_2\right)\frac{1}{|a_1|}dx_2$$

Functions of two or more *R.V.*s - 7

- **Ex75 (Book Ex3.9.6) Maximum and Minimum of Random Sample.** Suppose that $X_1,\ldots,X_n$ form a random sample of size n from a distribution for which the p.d.f. is f and the c.d.f. is F. The largest value Y_n and the smallest value Y_1 in the random sample are defined as:

$$Y_n = \max\{X_1,\ldots,X_n\}, Y_1 = \min\{X_1,\ldots,X_n\}.$$

$$G_n(y) = \Pr(Y_n \le y) = \Pr(X_1 \le y, X_2 \le y, \ldots, X_n \le y)$$
$$= \Pr(X_1 \le y)\Pr(X_2 \le y)\ldots\Pr(X_n \le y) = [F(y)]^n$$

$$g_n(y) = n[F(y)]^{n-1} f(y) \quad \text{for } -\infty < y < \infty.$$

$$G_1(y) = \Pr(Y_1 \le y) = 1 - \Pr(Y_1 > y)$$
$$= 1 - \Pr(X_1 > y, X_2 > y, \ldots, X_n > y)$$
$$= 1 - \Pr(X_1 > y)\Pr(X_2 > y)\ldots\Pr(X_n > y) = 1 - [1 - F(y)]^n.$$

$$g_1(y) = n[1 - F(y)]^{n-1} f(y) \quad \text{for } -\infty < y < \infty.$$

Functions of two or more *R.V.*s - 8

For more information of general functions of tow *R.V.*s, please see :

1.Textbool1 P182-P186.

2.Textbook2 P276-P278.