04. Discrete and Continuous Probability Distribution

1. Discrete Probability Distribution
2. Continuous Probability Distribution

Part 1

Discrete Probability Distribution

Adapted From :

Probability & Statistics for Engineers & Scientists, 9th Ed.

Walpole/Myers/Myers/Ye (c)2012

Probability & Statistics for Engineering & The Science, 9th Ed.

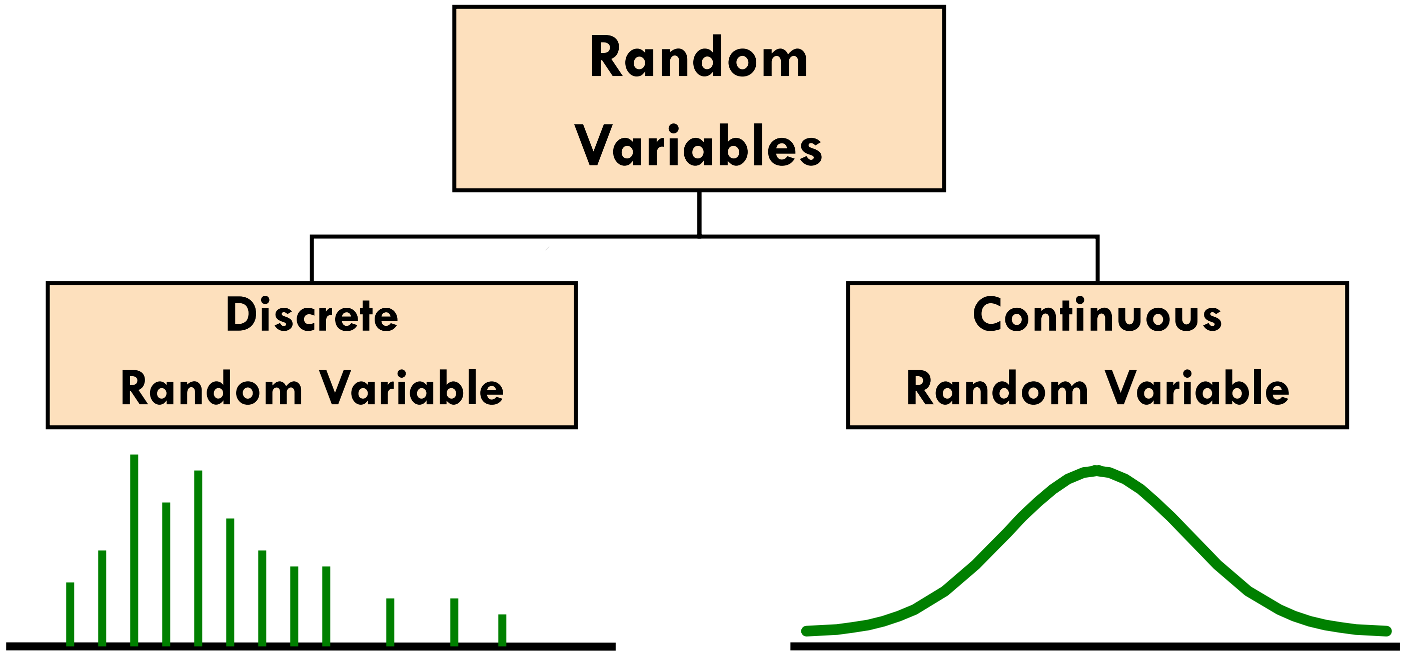
J.L.Devore © 2014

Statistics for Managers, 5th Ed

Levine/Stephan/Krehbiel/Berenson © 2008

Introduction to Probability Distributions

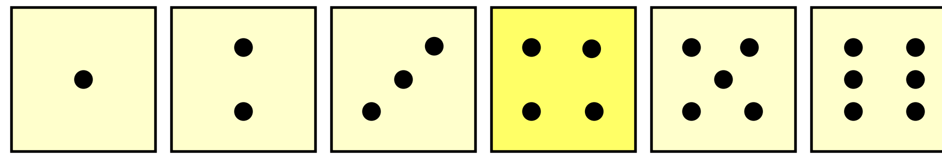
- **Random Variable**
  - Represents a possible numerical value from an uncertain event



Discrete Random Variables

- Can only assume a countable number of values

Examples:



- **Roll a die twice**

  Let X be the number of times 4 comes up

  (then X could be 0, 1, or 2 times)

- **Toss a coin 5 times.**

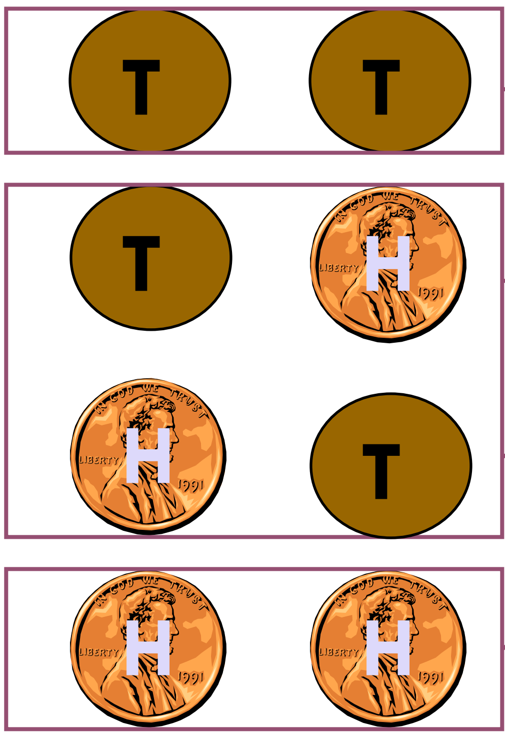
  Let X be the number of heads

  (then X = 0, 1, 2, 3, 4, or 5)

Discrete Probability Distribution

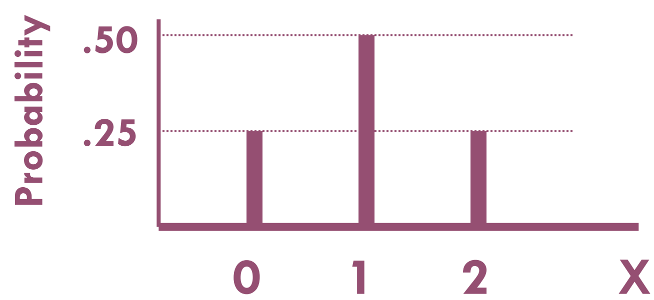
Experiment: Toss 2 Coins. Let X = # heads.

4 possible outcomes



Probability Distribution

X Value	Probability
0	1/4 = .25
1	2/4 = .50
2	1/4 = .25



Discrete Random Variable Summary Measures

- Expected Value of a discrete distribution

  (Weighted Average)

$$\mu = E(X) = \sum_{i=1}^{N} X_i P(X_i)$$

- Example: Toss 2 coins, X = # of heads, compute expected value of X:

E(X) = (0 x .25) + (1 x .50) + (2 x .25)

= 1.0

X	P(X)
0	.25
1	.50
2	.25

Discrete Random Variable Summary Measures

(continued)

- Variance of a discrete random variable

$$\sigma^2 = \sum_{i=1}^{N}[X_i - E(X)]^2 P(X_i)$$

- Standard Deviation of a discrete random variable

$$\sigma = \sqrt{\sigma^2} = \sqrt{\sum_{i=1}^{N}[X_i - E(X)]^2 P(X_i)}$$

where:

E(X) = Expected value of the discrete random variable X

X_i = the i^{th} outcome of X

$P(X_i)$ = Probability of the i^{th} occurrence of X

Probability Distributions

- Certain probability distributions occur over and over in the real world.
  - Probability tables are published, mean and standard deviation is calculated, to make applying them easier.
  - The key is to apply the correct distribution based on the characteristics of the problem being studied.

Discrete Probability Distributions

- Certain probability distributions occur over and over in the real world.
  - Probability tables are published, mean and standard deviation is calculated, to make applying them easier.
  - The key is to apply the correct distribution based on the characteristics of the problem being studied.
- Some common discrete distribution models:
- Uniform: All outcomes are equally likely.
- Binomial: Number of successes in n independent trials, with each trial having probability of success p and probability of failure q (= 1-p).
- Multinomial: # of outcomes in n trials, with each of k possible outcomes having probabilities $p_1, p_2, \ldots, p_k$.

Discrete Probability Distributions

- Common discrete distribution models, continued:
- Hypergeometric: A sample of size n is selected from N items without replacement, and k items are classified as successes (N - k are failures).
- Negative Binomial: In n independent trials, with probability of success p and probability of failure q
(q = 1-p) on each trial, the probability that the kth success occurs on the xth trial.
- Geometric: Special case of the negative binomial. The probability that the 1st success occurs on the xth trial.
- Poisson: If λ is the rate of occurrence of an event (number of outcomes per unit time), the probability that x outcomes occur in a time interval of length t.

Discrete Uniform Distribution

- When X assumes the values x_1, x_2, …, x_k and each outcome is equally likely. Then

$$f(x;k) = \frac{1}{k}, x = x_1, x_2, \ldots, x_k,$$

and

$$\mu = \frac{\sum_{i=1}^{k} x_i}{k} \qquad \sigma^2 = \frac{\sum_{i=1}^{k} (x_i - \mu)^2}{k}$$

Since all observations are equally likely, this is similar to the mean and variance of a sample of size k, but note that we use <u>k rather than k - 1</u> to calculate variance.

Binomial Distribution

- Binomial: Number of successes in n independent trials, with each trial having probability of success p and probability of failure q (= 1-p).
    - Each trial is called a Bernoulli trial.
    - Experiment consists of n repeated trials.
    - Two possible outcomes, called success or failure.
    - P(success) = p, constant from trial to trial.
    - Each trial is independent.
- The name binomial comes from the binomial expansion of $(p + q)^n$, which equals

$$\binom{n}{0}q^{n} + \binom{n}{1}q^{n-1}p + \binom{n}{2}q^{n-2}p^{2} + \ldots + \binom{n}{n}p^{n}$$

Binomial Distribution

- Binomial: If x is the number of successes in n trials, each with two outcomes where p is the probability of success and q = 1-p is the probability of failure, the probability distribution of X is

$$b(x;n,p) = \binom{n}{x} p^x q^{n-x}, x = 0,1,2,\ldots,n.$$

the number of ways a given outcome x can occur times the probability of that outcome occurring, and $\mu = np$ $\sigma^2 = npq$

Binomial Probability Distribution

Bernoulli Process :

- A fixed number of observations, n (repeated trials)
  - e.g.: 15 tosses of a coin; ten light bulbs taken from a shipment
- Each trial results in an outcome that may be classified as a success or failure
  - e.g.: head or tail in each toss of a coin; defective or not defective light bulb
  - Generally called “success” and “failure”
  - Probability of success is p, probability of failure is q
- Constant probability for each observation (p)
  - e.g.: Probability of getting a tail is the same each time we toss the coin
- Repeated trials are independent

Binomial Probability Distribution *(continued)*

- Observations are independent
  - The outcome of one observation does not affect the outcome of the other
- Two sampling methods
  - Infinite population without replacement
  - Finite population with replacement

Possible Binomial Distribution Settings

- A manufacturing plant labels items as either defective or acceptable
- A firm bidding for contracts will either get a contract or not
- A marketing research firm receives survey responses of “yes I will buy” or “no I will not”
- New job applicants either accept the offer or reject it

Rule of Combinations

- The number of combinations of selecting X objects out of n objects is

$$\binom{n}{X} = \frac{n!}{X!(n-X)!}$$

where:

$$n! = n(n-1)(n-2)\ldots(2)(1)$$

$$X! = X(X-1)(X-2)\ldots(2)(1)$$

$$0! = 1 \quad \text{(by definition)}$$

Binomial Distribution Formula

$$P(X) = \frac{n!}{X!(n-X)!} p^X (q)^{n-X}$$

P(X) = probability of **X** successes in **n** trials,
with probability of success **p** on each trial

X = number of 'successes' in sample,
(X = 0, 1, 2, ..., n)

n = sample size (number of trials
or observations)

p = probability of "success"

q = 1 - p

Example: Flip a coin four times, let x = # heads:

n = 4

p = 0.5

q = (1 - .5) = .5

X = 0, 1, 2, 3, 4

Example:
Calculating a Binomial Probability

What is the probability of one success in five observations if the probability of success is .1?

$X = 1$, $n = 5$, and $p = .1$

$$P(X=1) = \frac{n!}{X!(n-X)!} p^X (q)^{n-X}$$

$$= \frac{5!}{1!(5-1)!} (.1)^1 (1-.1)^{5-1}$$

$$= (5)(.1)(.9)^4$$

$$= .32805$$

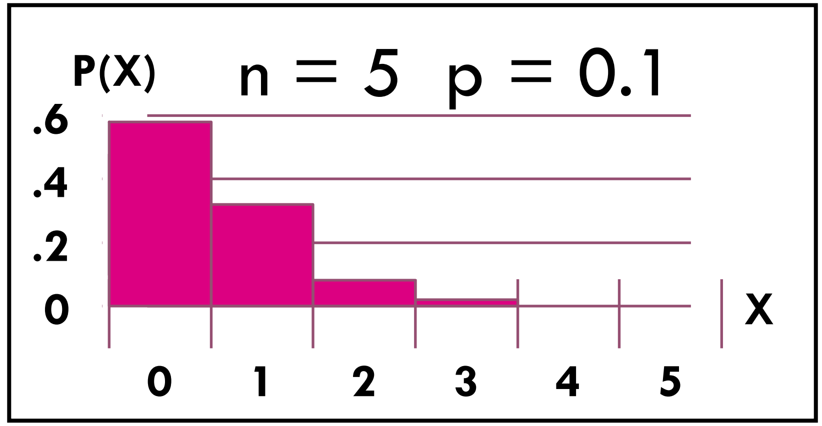
Binomial Probability Table

Table A.1 (continued) Binomial Probability Sums $\sum_{x=0}^{r} b(x; n, p)$

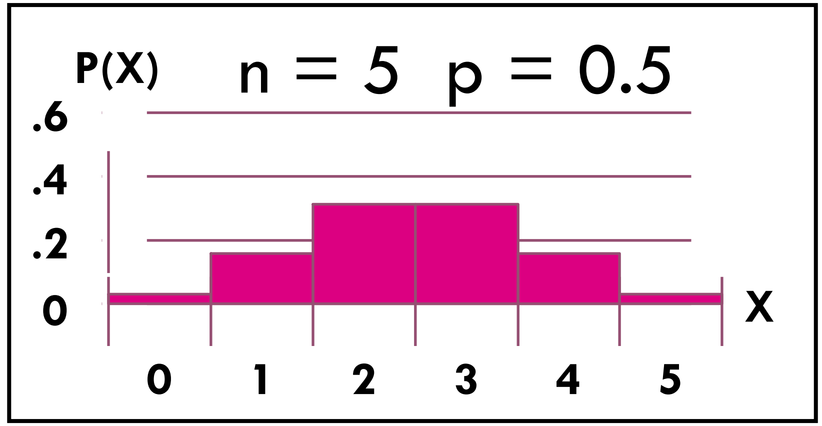
		p									
n	r	**0.10**	**0.20**	**0.25**	**0.30**	**0.40**	**0.50**	**0.60**	**0.70**	**0.80**	**0.90**
8	**0**	0.4305	0.1678	0.1001	0.0576	0.0168	0.0039	0.0007	0.0001	0.0000	
	1	0.8131	0.5033	0.3671	0.2553	0.1064	0.0352	0.0085	0.0013	0.0001	
	2	0.9619	0.7969	0.6785	0.5518	0.3154	0.1445	0.0498	0.0113	0.0012	0.0000
	3	0.9950	0.9437	0.8862	0.8059	0.5941	0.3633	0.1737	0.0580	0.0104	0.0004
	4	0.9996	0.9896	0.9727	0.9420	0.8263	0.6367	0.4059	0.1941	0.0563	0.0050
	5	1.0000	0.9988	0.9958	0.9887	0.9502	0.8555	0.6846	0.4482	0.2031	0.0381
	6		0.9999	0.9996	0.9987	0.9915	0.9648	0.8936	0.7447	0.4967	0.1869
	7		1.0000	1.0000	0.9999	0.9993	0.9961	0.9832	0.9424	0.8322	0.5695
	8				1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
9	**0**	0.3874	0.1342	0.0751	0.0404	0.0101	0.0020	0.0003	0.0000		
	1	0.7748	0.4362	0.3003	0.1960	0.0705	0.0195	0.0038	0.0004	0.0000	
	2	0.9470	0.7382	0.6007	0.4628	0.2318	0.0898	0.0250	0.0043	0.0003	0.0000
	3	0.9917	0.9144	0.8343	0.7297	0.4826	0.2539	0.0994	0.0253	0.0031	0.0001
	4	0.9991	0.9804	0.9511	0.9012	0.7334	0.5000	0.2666	0.0988	0.0196	0.0009
	5	0.9999	0.9969	0.9900	0.9747	0.9006	0.7461	0.5174	0.2703	0.0856	0.0083
	6	1.0000	0.9997	0.9987	0.9957	0.9750	0.9102	0.7682	0.5372	0.2618	0.0530
	7		1.0000	0.9999	0.9996	0.9962	0.9805	0.9295	0.8040	0.5638	0.2252
	8			1.0000	1.0000	0.9997	0.9980	0.9899	0.9596	0.8658	0.6126
	9					1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
10	**0**	0.3487	0.1074	0.0563	0.0282	0.0060	0.0010	0.0001	0.0000		
	1	0.7361	0.3758	0.2440	0.1493	0.0464	0.0107	0.0017	0.0001	0.0000	
	2	0.9298	0.6778	0.5256	0.3828	0.1673	0.0547	0.0123	0.0016	0.0001	
	3	0.9872	0.8791	0.7759	0.6496	0.3823	0.1719	0.0548	0.0106	0.0009	0.0000
	4	0.9984	0.9672	0.9219	0.8497	0.6331	0.3770	0.1662	0.0473	0.0064	0.0001

Binomial Distribution

- The shape of the binomial distribution depends on the values of p and n



- Here, n = 5 and p = .1



- Here, n = 5 and p = .5

Binomial Distribution Characteristics

- Mean

$$\mu = E(x) = np$$

- Variance and Standard Deviation

$$\sigma^2 = np(1-p) = npq$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{npq}$$

Where n = sample size
p = probability of success
(1 – p) = probability of failure

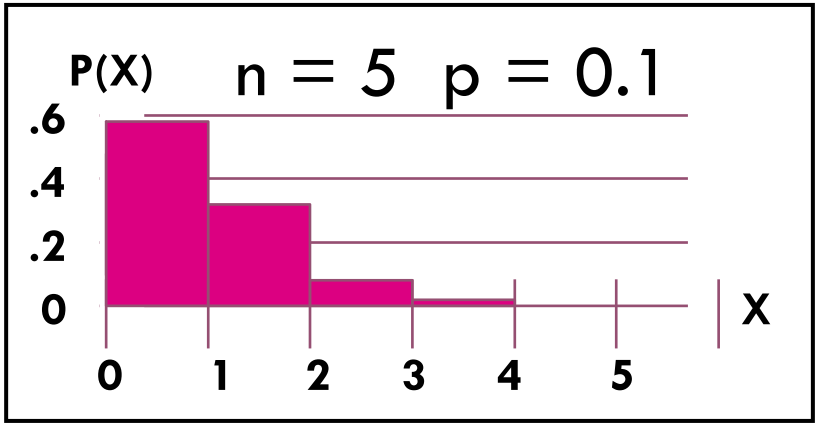
Binomial Characteristics

Examples

$$\mu = np = (5)(.1) = 0.5$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{(5)(.1)(1-.1)}$$

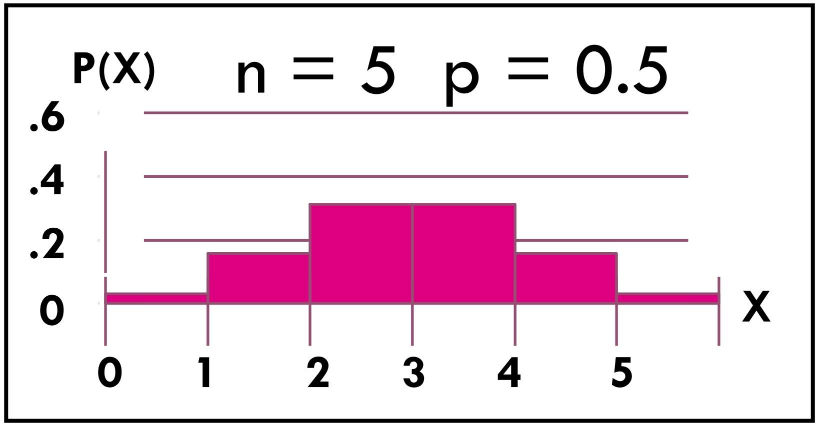
$$= 0.6708$$



$$\mu = np = (5)(.5) = 2.5$$

$$\sigma = \sqrt{np(1-p)} = \sqrt{(5)(.5)(1-.5)}$$

$$= 1.118$$



Using Binomial Tables

n = 10									
x	…	p=.20	p=.25	p=.30	**p=.35**	p=.40	p=.45	p=.50	
0	…	0.1074	0.0563	0.0282	0.0135	0.0060	0.0025	0.0010	10
1	…	0.2684	0.1877	0.1211	0.0725	0.0403	0.0207	0.0098	9
2	…	0.3020	0.2816	0.2335	0.1757	0.1209	0.0763	0.0439	8
3	…	0.2013	0.2503	0.2668	**0.2522**	0.2150	0.1665	0.1172	7
4	…	0.0881	0.1460	0.2001	0.2377	0.2508	0.2384	0.2051	6
5	…	0.0264	0.0584	0.1029	0.1536	0.2007	0.2340	0.2461	5
6	…	0.0055	0.0162	0.0368	0.0689	0.1115	0.1596	0.2051	4
7	…	0.0008	0.0031	0.0090	0.0212	0.0425	0.0746	0.1172	3
8	…	0.0001	**0.0004**	0.0014	0.0043	0.0106	0.0229	0.0439	**2**
9	…	0.0000	0.0000	0.0001	0.0005	0.0016	0.0042	0.0098	1
10	…	0.0000	0.0000	0.0000	0.0000	0.0001	0.0003	0.0010	0
	…	p=.80	**p=.75**	p=.70	p=.65	p=.60	p=.55	p=.50	**x**

Examples:

n = 10, p = .35, x = 3: b(3;10,.35) = .2522

n = 10, p = .75, x = 2: b(2;10,.75) = .0004

Binomial Distribution Examples

1. Sebuah perusahaan farmasi memperhatikan 5 orang pekerjanya yang sering terlambat hadir. Peluang hadir terlambat dari setiap pekerja adalah 0,4 dan mereka hadir tidak tergantung satu dengan lainnya. Berapa peluang 2 orang pekerja dari 5 pekerja hadir terlambat ?

   Jawab : $P(2) = \dfrac{5\,!}{2\,!\,(5-2)\,!}(0{,}4)^2\ (0{,}6)^3$

   $= (120/12)\ (0{,}16)\,(0{,}216)$

   $= 0{,}3456$

2. Pemilik toko elektronik memperhatikan bahwa peluang seorang pengunjung datang membeli adalah 0,3. Jika pada suatu hari ada 15 org pengunjung, berapa :
   a. Peluang paling sedikit 1 orang pengunjung yang membeli ?
   b. Peluang tidak lebih dari 4 pengunjung yang membeli ?

   Jawab : a. $P(x \geq 1) = 1 - P(0) = 1 - 0{,}0047 = 0{,}9953$

   b. $P(x \leq 4) = P(0) + P(1) + P(2) + P(3) + P(4) = 0{,}5155$

3. Harley Davidson, seorang supervisor pengendalian mutu PT Kyoto melakukan tugas rutinnya mengecek transmisi otomatis. Prosedurnya, Ia memindahkan 10 unit transmisi dan memeriksa kerusakannya. Data masa lalu menunjukkan 2% dari transmisi mengalami kerusakan.
   a. Berapa peluang lebih dari 2 transmisi pada sampel yang mengalami kerusakan ? (Ans. 0,0009)
   b. Berapa peluang tidak ada transmisi yang rusak ? (Ans. 0,8171)

Multinomial Distribution

- Multinomial: Number of outcomes in n trials, with each of k possible outcomes for each trial having probabilities $p_1, p_2, \ldots, p_k$.
  - Generalization of binomial to k outcomes.
  - Experiment consists of n repeated trials.
  - P(outcome i) = p_i, constant from trial to trial.
  - Repeated trials are independent.

- As with the binomial, the name multinomial comes from the multinomial expansion of $(p_1 + p_2 + \ldots + p_k)^n$.

Multinomial Distribution

- <u>Multinomial</u>: The distribution of the number, x_i, of each of k types of outcomes in n trials, with the k outcomes having probabilities p_1, p_2, …, p_k, is

$$f(x_1, x_2, \ldots, x_k; p_1, p_2, \ldots, p_k, n) = \binom{n}{x_1, x_2, \ldots, x_k} p_1^{x_1} p_2^{x_2} \ldots p_k^{x_k}$$

with $\sum_{i=1}^{k} x_i = n$

and $\sum_{i=1}^{k} p_i = 1$

What about μ and σ^2?

- μ and σ only apply to distributions of a single random variable

Example 5.7

Example 5.7: The complexity of arrivals and departures of planes at an airport is such that computer simulation is often used to model the "ideal" conditions. For a certain airport with three runways, it is known that in the ideal setting the following are the probabilities that the individual runways are accessed by a randomly arriving commercial jet:

Runway 1: $p_1 = 2/9$,
Runway 2: $p_2 = 1/6$,
Runway 3: $p_3 = 11/18$.

What is the probability that 6 randomly arriving airplanes are distributed in the following fashion?

Runway 1: 2 airplanes,
Runway 2: 1 airplane,
Runway 3: 3 airplanes

***Solution*:** Using the multinomial distribution, we have

$$f\left(2, 1, 3; \frac{2}{9}, \frac{1}{6}, \frac{11}{18}, 6\right) = \binom{6}{2, 1, 3}\left(\frac{2}{9}\right)^2\left(\frac{1}{6}\right)^1\left(\frac{11}{18}\right)^3$$
$$= \frac{6!}{2!\,1!\,3!} \cdot \frac{2^2}{9^2} \cdot \frac{1}{6} \cdot \frac{11^3}{18^3} = 0.1127.$$

Hypergeometric Distribution

- Hypergeometric: The distribution of the number of successes, x, in a sample of size n is selected from N items without replacement, where k items are classified as successes (and N - k as failures), is

$$h(x;N,n,k)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}},\max\{0,n-(N-k)\}\leq x\leq \min\{k,n\}$$

then $\mu=\frac{nk}{N}$

and $\sigma^2=\frac{N-n}{N-1}\cdot n\cdot\frac{k}{N}(1-\frac{k}{N})$

Binomial Approximation to Hypergeometric

- If n is small compared with N, then the hypergeometric distribution can be approximated using the binomial distribution.
- The rule of thumb is that this is valid if $(n/N) \leq 0.05$. In this case, we can use the binomial distribution with parameters n and $p = k/N$.
- Then $\mu = np = \frac{nk}{N}$ $\sigma^2 = npq = n \cdot \frac{k}{N}(1 - \frac{k}{N})$

The Hypergeometric Distribution

- “n” trials in a sample taken from a finite population of size N
- Sample taken without replacement
- Trials are dependent
- Concerned with finding the probability of “X” successes in the sample where there are “k” successes in the population

Hypergeometric Distribution Formula

(Two possible outcomes per trial)

$$P(X) = h(x; N, n, k) = \frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}$$

Where

N = Population size

k = number of successes in the population

N – k = number of failures in the population

n = sample size

x = number of successes in the sample

n – x = number of failures in the sample

Properties of the Hypergeometric Distribution

- The mean of the hypergeometric distribution is

$$\boxed{\mu = \mathrm{E}(\mathrm{x}) = \frac{\mathrm{nk}}{N}}$$

- The variance is

$$\boxed{\sigma^2 = \frac{N-n}{N-1}.n.\frac{k}{N}\left(1-\frac{k}{N}\right)}$$

Using the Hypergeometric Distribution

- Example: 3 different computers are checked from 10 in the department. 4 of the 10 computers have illegal software loaded. What is the probability that 2 of the 3 selected computers have illegal software loaded?

N = 10	n = 3
k = 4	x = 2

$$P(X=2)=\frac{\binom{k}{x}\binom{N-k}{n-x}}{\binom{N}{n}}=\frac{\binom{4}{2}\binom{6}{1}}{\binom{10}{3}}=\frac{(6)(6)}{120}=0.3$$

The probability that 2 of the 3 selected computers will have illegal software loaded is .30, or 30%.

Negative Binomial Distribution

- <u>Negative Binomial</u>: In n independent trials, with probability of success p and probability of failure q
  (q = 1-p) on each trial, the probability that the kth success occurs on the xth trial.

$$b^*(x;k,p) = \binom{x-1}{k-1} p^k q^{x-k}, x = k, k+1, k+2, \ldots$$

  Again we have the number of ways an outcome x can occur times the probability of that outcome occurring.

- The above formula comes from the fact that in order to to get the kth success on the xth trial, we must have
  k - 1 successes in the first x - 1 trials, and then the final trial must also be a success.

Example

Example 5.14: In an NBA (National Basketball Association) championship series, the team that wins four games out of seven is the winner. Suppose that teams A and B face each other in the championship games and that team A has probability 0.55 of winning a game over team B.

(a) What is the probability that team A will win the series in 6 games?

(b) What is the probability that team A will win the series?

(c) If teams A and B were facing each other in a regional playoff series, which is decided by winning three out of five games, what is the probability that team A would win the series?

Solution: (a) $b^*(6; 4, 0.55) = \binom{5}{3} 0.55^4 (1 - 0.55)^{6-4} = 0.1853$

(b) P(team A wins the championship series) is

$$\begin{aligned} &b^*(4; 4, 0.55) + b^*(5; 4, 0.55) + b^*(6; 4, 0.55) + b^*(7; 4, 0.55) \\ &= 0.0915 + 0.1647 + 0.1853 + 0.1668 = 0.6083. \end{aligned}$$

(c) P(team A wins the playoff) is

$$\begin{aligned} &b^*(3; 3, 0.55) + b^*(4; 3, 0.55) + b^*(5; 3, 0.55) \\ &= 0.1664 + 0.2246 + 0.2021 = 0.5931. \end{aligned}$$

Geometric Distribution

- Geometric: Special case of the negative binomial with k = 1. The probability that the 1st success occurs on the xth trial is

$$g(x; p) = pq^{x-1}, x = 1,2,3,\ldots$$

then

$$\mu = \frac{1}{p}$$

and

$$\sigma^2 = \frac{1-p}{p^2}$$

Example 5.15

- For a certain manufacturing process, it is known that, on the average, 1 in every 100 items is defective. What is the probability that the fifth item inspected is the first defective item found?
- Solution :

Using the geometric distribution with $x = 5$ and $p = 0.01$, we have:

$$g(5; 0.01) = (0.01)(0.99)4 = 0.0096.$$

Example 3.12

Starting at a fixed time, **we observe the gender of each newborn child at a certain hospital until a boy (*B*) is born.**

Let $p = P(B)$, assume that successive births are independent, and define the rv X by x = number of births observed.

Then

$$
\begin{aligned}
p(1) &= P(X = 1) \\
&= P(B) \\
&= p
\end{aligned}
$$

Example 3.12

cont'd

$$p(2) = P(X = 2)$$
$$= P(GB)$$
$$= P(G) \cdot P(B)$$
$$= (1 - p)p$$

and

$$p(3) = P(X = 3)$$
$$= P(GGB)$$
$$= P(G) \cdot P(G) \cdot P(B)$$
$$= (1 - p)^2 p$$

Poisson Distribution

- Poisson distribution: If λ is the average # of outcomes per unit time (arrival rate), the Poisson distribution gives the probability that x outcomes occur in a given time interval of length t.
- A Poisson process has the following properties:
  - Memoryless: the number of occurrences in one time interval is independent of the number in any other disjoint time interval.
  - The probability that a single outcome will occur during a very short time interval is proportional to the size of the time interval and independent of other intervals.
  - The probability that more than one outcome will occur in a very short time interval is negligible.
- Note that the rate could be per unit length, area, or volume, rather than time.

Poisson Distribution

- <u>Poisson distribution</u>: If λ is the rate of occurrence of an event (average # of outcomes per unit time), the probability that x outcomes occur in a time interval of length t is

$$p(x;\lambda t) = \frac{e^{-\lambda t}(\lambda t)^{x}}{x!}, x = 0,1,2,...$$

then

$$\mu = \sigma^{2} = \lambda t$$

Poisson Distribution Characteristics

- Mean

$$\mu = \lambda t$$

- Variance and Standard Deviation

$$\sigma^2 = \lambda t$$

$$\sigma = \sqrt{(\lambda t)}$$

Poisson Probability Sums

Table A.2 Poisson Probability Sums $\sum_{x=0}^{r} p(x;\mu)$

	μ								
r	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**	**0.6**	**0.7**	**0.8**	**0.9**
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.9953	0.9825	0.9631	0.9384	0.9098	0.8781	0.8442	0.8088	0.7725
2	0.9998	0.9989	0.9964	0.9921	0.9856	0.9769	0.9659	0.9526	0.9371
3	1.0000	0.9999	0.9997	0.9992	0.9982	0.9966	0.9942	0.9909	0.9865
4		1.0000	1.0000	0.9999	0.9998	0.9996	0.9992	0.9986	0.9977
5				1.0000	1.0000	1.0000	0.9999	0.9998	0.9997
6							1.0000	1.0000	1.0000

	μ								
r	**1.0**	**1.5**	**2.0**	**2.5**	**3.0**	**3.5**	**4.0**	**4.5**	**5.0**
0	0.3679	0.2231	0.1353	0.0821	0.0498	0.0302	0.0183	0.0111	0.0067
1	0.7358	0.5578	0.4060	0.2873	0.1991	0.1359	0.0916	0.0611	0.0404
2	0.9197	0.8088	0.6767	0.5438	0.4232	0.3208	0.2381	0.1736	0.1247
3	0.9810	0.9344	0.8571	0.7576	0.6472	0.5366	0.4335	0.3423	0.2650
4	0.9963	0.9814	0.9473	0.8912	0.8153	0.7254	0.6288	0.5321	0.4405
5	0.9994	0.9955	0.9834	0.9580	0.9161	0.8576	0.7851	0.7029	0.6160
6	0.9999	0.9991	0.9955	0.9858	0.9665	0.9347	0.8893	0.8311	0.7622
7	1.0000	0.9998	0.9989	0.9958	0.9881	0.9733	0.9489	0.9134	0.8666
8		1.0000	0.9998	0.9989	0.9962	0.9901	0.9786	0.9597	0.9319
9			1.0000	0.9997	0.9989	0.9967	0.9919	0.9829	0.9682
10				0.9999	0.9997	0.9990	0.9972	0.9933	0.9863

Using Poisson Tables

X	λ								
	0.10	0.20	0.30	0.40	**0.50**	0.60	0.70	0.80	0.90
0	0.9048	0.8187	0.7408	0.6703	0.6065	0.5488	0.4966	0.4493	0.4066
1	0.0905	0.1637	0.2222	0.2681	0.3033	0.3293	0.3476	0.3595	0.3659
2	0.0045	0.0164	0.0333	0.0536	**0.0758**	0.0988	0.1217	0.1438	0.1647
3	0.0002	0.0011	0.0033	0.0072	0.0126	0.0198	0.0284	0.0383	0.0494
4	0.0000	0.0001	0.0003	0.0007	0.0016	0.0030	0.0050	0.0077	0.0111
5	0.0000	0.0000	0.0000	0.0001	0.0002	0.0004	0.0007	0.0012	0.0020
6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0002	0.0003
7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

Example: Find P(X = 2) if $\lambda t = .50$

$$P(X=2) = \frac{e^{-\lambda t}\lambda t^{X}}{X!} = \frac{\mathrm{e}^{-0.50}(0.50)^2}{2!} = .0758$$

The Poisson Distribution

- Properties of Poisson Distribution :
  - The number of outcomes occurring in one time interval or specified region is independent of the number that occurs in any other disjoint time interval or region of space
  - The probability that a single outcome will occur during a very short time interval or in a small regions is proportional to the length of the time interval or the size of the region and does not depend on the number of outcomes occurring outside this time interval or region.
  - The probability that more than one outcome will occur in such a short time interval or fall in such a small region is negligible.
  - The average number of outcomes per unit is λ (lambda)

Example :

\# Customers arriving in 15 min

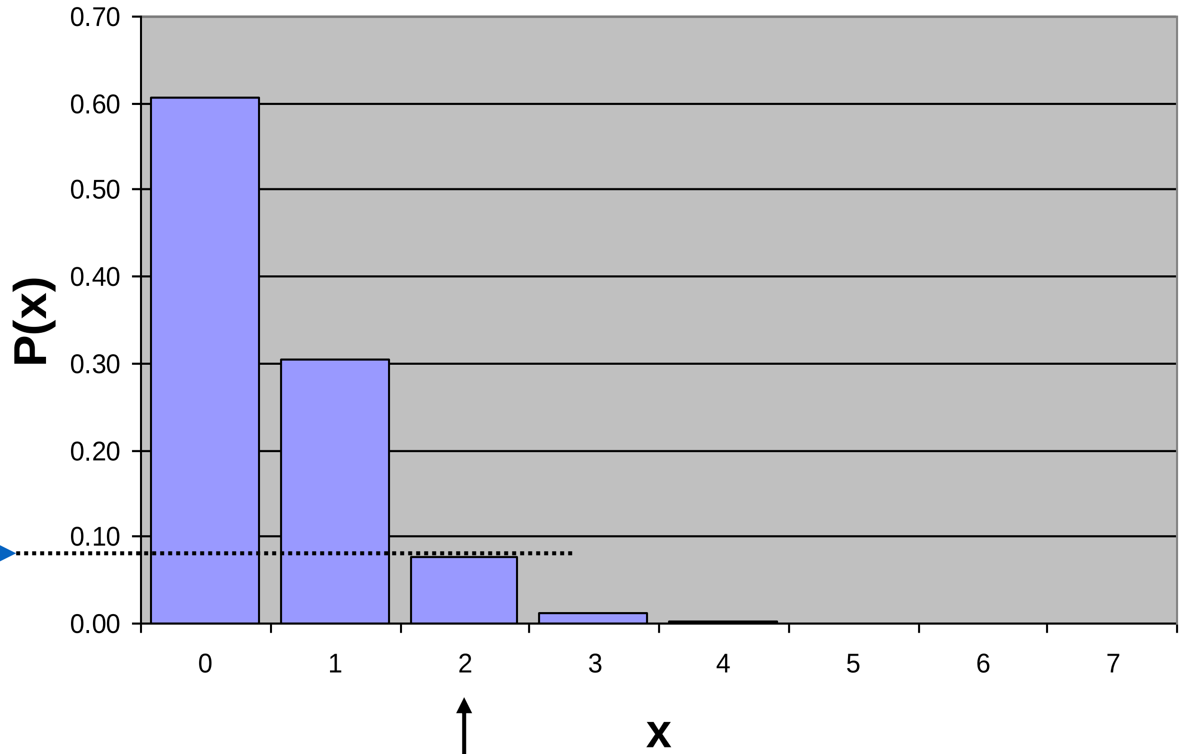
\# Defects per case of light bulbs

Graph of Poisson Probabilities

Graphically:

$\lambda = .50$

X	$\lambda =$ 0.50
0	0.6065
1	0.3033
2	0.0758
3	0.0126
4	0.0016
5	0.0002
6	0.0000
7	0.0000

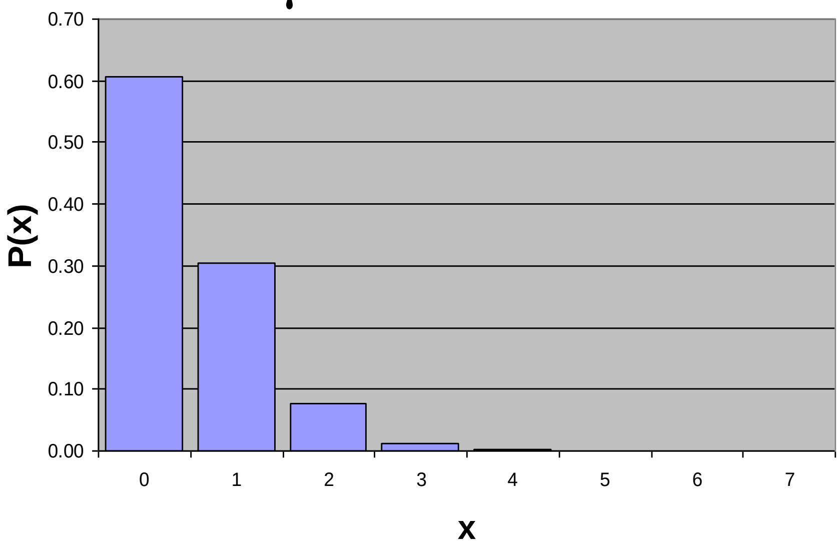


P(X = 2) = .0758

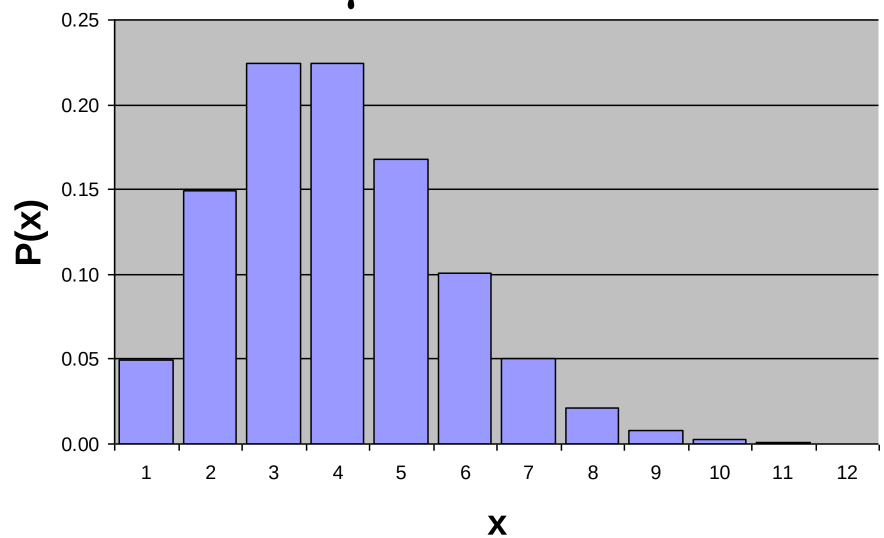
Poisson Distribution Shape

- The shape of the Poisson Distribution depends on the parameter μ :

μ = 0.50



μ = 3.00



Example 5.17 and 5.18

Example 5.17: During a laboratory experiment, the average number of radioactive particles passing through a counter in 1 millisecond is 4. What is the probability that 6 particles enter the counter in a given millisecond?

***Solution*:** Using the Poisson distribution with $x = 6$ and $\lambda t = 4$ and referring to Table A.2, we have

$$p(6;4) = \frac{e^{-4}4^6}{6!} = \sum_{x=0}^{6} p(x;4) - \sum_{x=0}^{5} p(x;4) = 0.8893 - 0.7851 = 0.1042.$$

Example 5.18: Ten is the average number of oil tankers arriving each day at a certain port. The facilities at the port can handle at most 15 tankers per day. What is the probability that on a given day tankers have to be turned away?

***Solution*:** Let X be the number of tankers arriving each day. Then, using Table A.2, we have

$$P(X > 15) = 1 - P(X \leq 15) = 1 - \sum_{x=0}^{15} p(x;10) = 1 - 0.9513 = 0.0487.$$

Like the binomial distribution, the Poisson distribution is used for quality control, quality assurance, and acceptance sampling. In addition, certain important continuous distributions used in reliability theory and queuing theory depend on the Poisson process. Some of these distributions are discussed and developed in Chapter 6. The following theorem concerning the Poisson random variable is given in Appendix A.25.

Poisson Distribution Examples

1. Setiap minggu rata-rata 5 ekor burung merpati mati karena menabrak Monumen di taman kota. Kelompok pecinta lingkungan hidup minta PEMDA setempat untuk mengalokasikan dana guna mencegah terjadi hal tersebut. PEMDA menyetujui bila peluang lebih dari 3 burung mati melebihi 0,7. Akankah PEMDA memberikan dana tersebut ?

   Jawab : $P(x > 3) = 1 - P(x \leq 3) = 1 - P(0) - P(1) - P(2) - P(3)$
   $= 0, 735$
   Karena peluangnya lebih besar dari 0,7 maka PEMDA memberikan dana

2. Seorang pianis merasa terganggu bila akan memulai konser, ia mendengar suara batuk dari penonton. Pada konser terakhirnya, ia menghitung ada 8 kali suara batuk terdengar saat konser akan dimulai. Pianis ini mengancam panitia konser bila malam ini terdengar lebih dari 5 kali suara batuk, ia akan membatalkan konser. Berapa peluang ia akan ikut konser malam ini ?

   Jawab : $P(x \leq 5) = P(0) + P(1) + P(2) + P(3) + P(4) + P(5)$
   $= 0{,}1912$

3. Sebuah mesin fotokopi mengalami gangguan 1 kali per 100 halaman kopian. Jika seorang harus memfotokopi 500 halaman . Berapa peluang terjadinya gangguan nol ?

   Jawab : untuk interval 100 halaman $\lambda t = 1$; untuk 500 hal. $\lambda t = 5$
   Maka $P(x=0) = 0{,}0067$

4. Dalam jam-jam sibuk nasabah yang datang 90 orang/jam. Berapa peluang 15 orang atau lebih nasabah yang datang selama 6 menit pada jam-jam sibuk ? (Ans. 0,0414)

Poisson Approximation to Binomial

- For a set of Bernoulli trials with n very large and p small, the Poisson distribution with mean np can be used to approximate the binomial distribution.
  - Needed since binomial tables only go up to n = 20.

The rule of thumb is that this approximation is valid if $n \geq 20$ and $p \leq 0.05$. (If $n \geq 100$, the approximation is excellent if $np \leq 10$). In this case, we can use the Poisson distribution with

$$\mu = \sigma^2 = np$$

A different approximation for the binomial can be used for large n if p is not small.

Example 5.19

- In a certain industrial facility, accidents occur infrequently. It is known that the probability of an accident on any given day is 0.005 and accidents are independent of each other.
  (a) What is the probability that in any given period of 400 days there will be an accident on one day?
  (b) What is the probability that there are at most three days with an accident?

***Solution*:** Let X be a binomial random variable with $n = 400$ and $p = 0.005$. Thus, $np = 2$. Using the Poisson approximation,

(a) $P(X = 1) = e^{-2}2^1 = 0.271$ and

(b) $P(X \leq 3) = \sum_{x=0}^{3} e^{-2}2^x/x! = 0.857$.

Example 5.20

- In a manufacturing process where glass products are made, defects or bubbles occur, occasionally rendering the piece undesirable for marketing. It is known that, on average, 1 in every 1000 of these items produced has one or more bubbles. What is the probability that a random sample of 8000 will yield fewer than 7 items possessing bubbles?

Solution: This is essentially a binomial experiment with $n = 8000$ and $p = 0.001$. Since p is very close to 0 and n is quite large, we shall approximate with the Poisson distribution using

$$\mu = (8000)(0.001) = 8.$$

Hence, if X represents the number of bubbles, we have

$$P(X < 7) = \sum_{x=0}^{6} b(x; 8000, 0.001) \approx p(x; 8) = 0.3134.$$

Summary :

Discrete Probability Distributions

- Binomial: Number of successes in n independent trials, with each trial having probability of success p and probability of failure q (= 1-p).
- Hypergeometric: A sample of size n is selected from N items without replacement, and k items are classified as successes (N - k are failures).
- Poisson: If λ is the rate of occurrence of an event (number of outcomes per unit time), the probability that x outcomes occur in a time interval of length t.

Part 2

Continuous Probability Distribution

Adapted From :

Probability & Statistics for Engineers & Scientists, 9th Ed.

Walpole/Myers/Myers/Ye (c)2012

Probability & Statistics for Engineering & The Science, 9th Ed.

J.L.Devore © 2014

Statistics for Managers, 5th Ed

Levine/Stephan/Krehbiel/Berenson © 2008

Continuous Probability Distributions

- A continuous random variable is a variable that can assume any value on a continuum (can assume an infinite number of values)
  - thickness of an item
  - time required to complete a task
  - temperature of a solution
  - height, in inches
- These can potentially take on any value, depending only on the ability to measure accurately.

Continuous Uniform Distribution

- The uniform distribution is a probability distribution that has equal probabilities for all possible outcomes of the random variable

- Also called a rectangular distribution

Continuous Uniform Distribution

- The continuous uniform distribution is the simplest of the continuous distributions. It models the situation where the probability is uniform in the interval [A,B].
- The density function is

$$
\begin{aligned}
f(x; A, B) &= \frac{1}{B-A}, \quad A \le x \le B \\
&= \quad 0 \qquad \text{elsewhere}.
\end{aligned}
$$

- This density function is a rectangle with base B - A. The height of the rectangle is?
  - The height is 1 / (B - A).

Uniform Distribution

- The mean and variance of the uniform distribution are

$$\mu = \frac{A+B}{2}$$

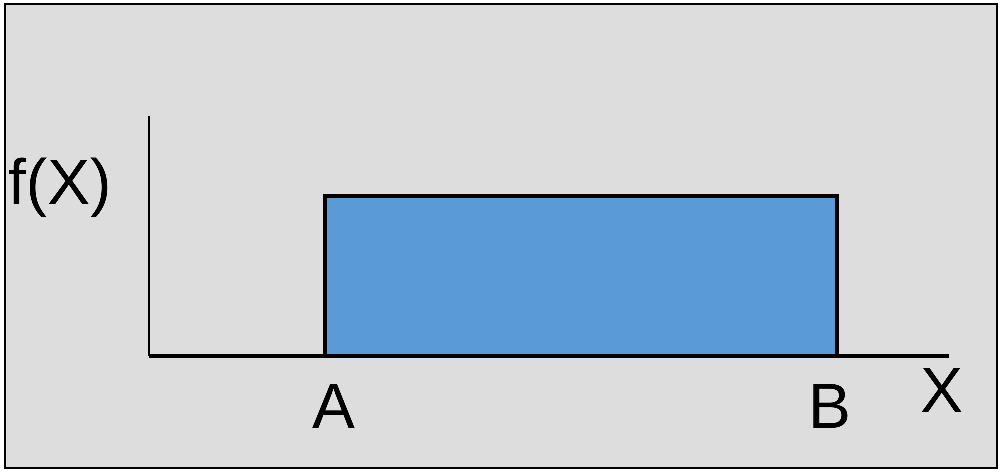
and

$$\sigma^2 = \frac{(B-A)^2}{12}$$

The Uniform Distribution

(continued)

The Continuous Uniform Distribution:



$$\mu = \frac{A+B}{2}$$

$$\sigma^2 = \frac{(B-A)^2}{12}$$

where

f(X) = value of the density function at any X value

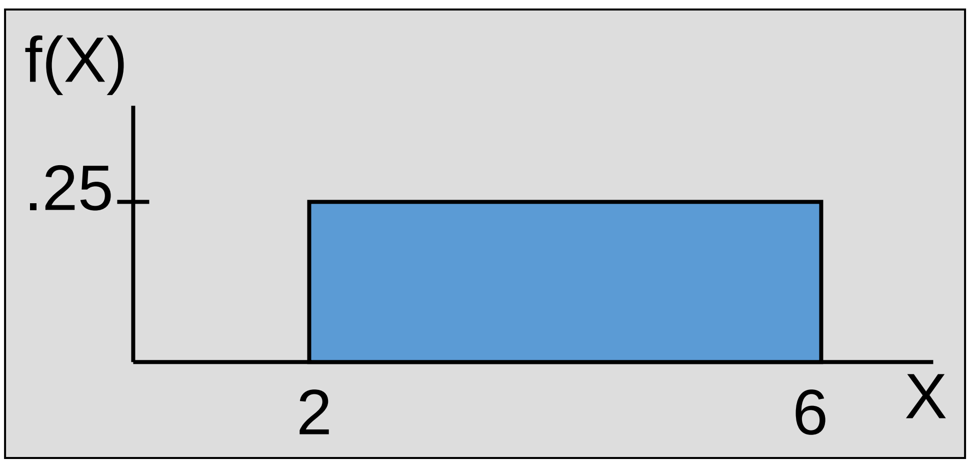
A = minimum value of X

B = maximum value of X

Uniform Distribution Example

Example: Uniform Probability Distribution
Over the range $2 \le X \le 6$:

$$f(X) = \frac{1}{6-2} = .25 \quad \text{for} \quad 2 \le X \le 6$$



$$\mu = \frac{A+B}{2} = \frac{2+6}{2} = 4$$

$$\sigma = \sqrt{\frac{(B-A)^2}{12}} = \sqrt{\frac{(6-2)^2}{12}} = 1.1547$$

Example 6.1

Suppose that a large conference room at a certain company can be reserved for no more than 4 hours. Both long and short conferences occur quite often. In fact, it can be assumed that the length X of a conference has a uniform distribution on the interval $[0, 4]$.

(a) What is the probability density function?

(b) What is the probability that any given conference lasts at least 3 hours?

Solution

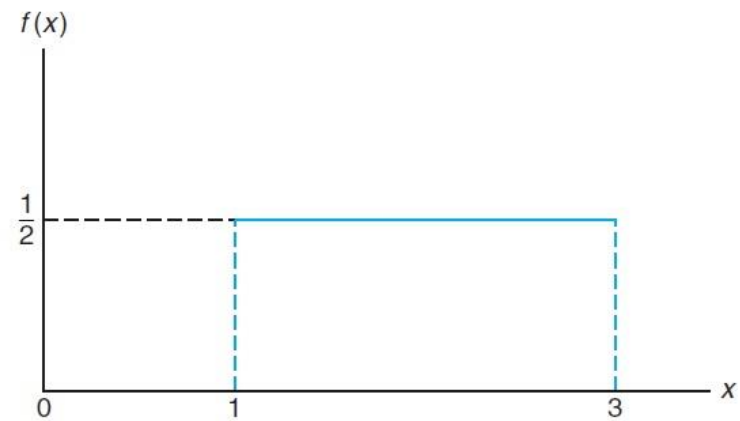


Figure 6.1: The density function for a random variable on the interval $[1, 3]$.

Solution: (a) The appropriate density function for the uniformly distributed random variable X in this situation is

$$f(x) = \begin{cases} \frac{1}{4}, & 0 \leq x \leq 4, \\ 0, & \text{elsewhere.} \end{cases}$$

(b) $P[X \geq 3] = \int_3^4 \frac{1}{4}\, dx = \frac{1}{4}$.

Example 4.4

- The direction of an imperfection with respect to a reference line on a circular object such as a tire, brake rotor, or flywheel is, in general, subject to uncertainty.

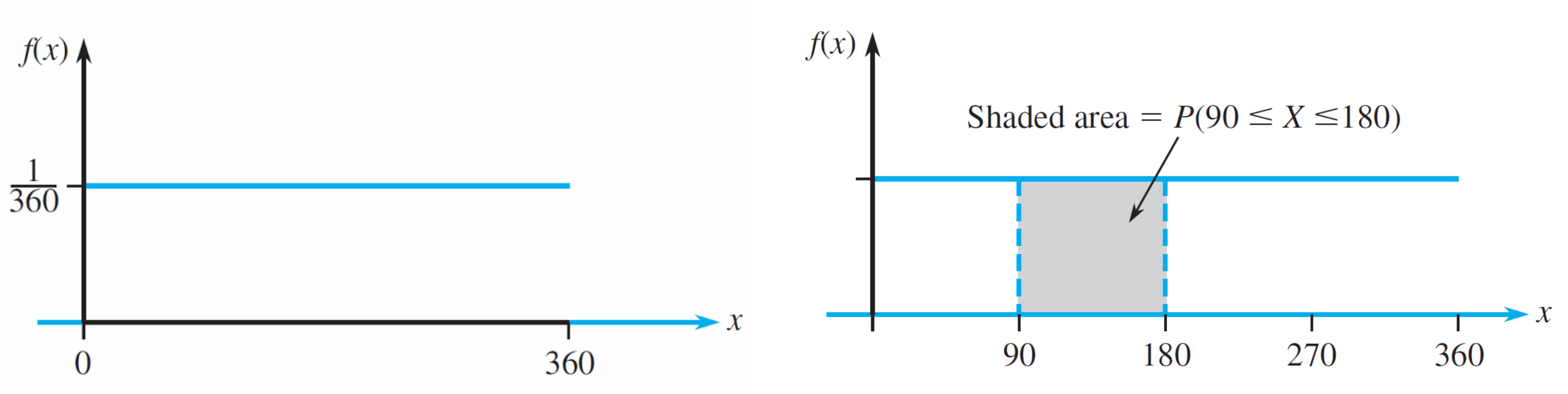
- Consider the reference line connecting the valve stem on a tire to the center point, and let *X* be the angle measured clockwise to the location of an imperfection. One possible pdf for *X* is

$$f(x) = \begin{cases} \dfrac{1}{360} & 0 \le x < 360 \\ 0 & \text{otherwise} \end{cases}$$

Example 4.4

cont'd

- The pdf is graphed in Figure 4.3.



The pdf and probability from Example 4

Figure 4.3

Example 4.4

cont'd

- Clearly $f(x) \geq 0$. The area under the density curve
- is just the area of a rectangle:

$$(\text{height})(\text{base}) = \left(\frac{1}{360}\right)(360) = 1.$$

- The probability that the angle is between 90° and 180° is

$$P(90 \leq X \leq 180) = \int_{90}^{180} \frac{1}{360}\, dx$$

$$= \frac{x}{360} \Big|_{x=90}^{x=180}$$

$$= \frac{1}{4} = .25$$

cont'd

Example 4.4

- The probability that the angle of occurrence is within 90° of the reference line is

- $$P(0 \leq X \leq 90) + P(270 \leq X < 360) = .25 + .25 = .50$$

Probability Distributions for Continuous Variables

- Because whenever $0 \le a \le b \le 360$ in Example 4.4 and $P(a \le X \le b)$ depends only on the width $b - a$ of the interval, X is said to have a uniform distribution.
- **Definition**

A continuous rv X is said to have a **uniform distribution** on the interval $[A, B]$ if the pdf of X is

$$f(x; A, B) = \begin{cases} \dfrac{1}{B - A} & A \le x \le B \\ 0 & \text{otherwise} \end{cases}$$

Normal Distribution

- The normal distribution is the most important and widely used distribution in statistics. Some examples:
  - This bell-shaped curve, sometimes called the Gaussian distribution, explains many natural phenomena.
  - Average age of the world's population.
  - Physical measurements like blood pressure.
  - Standardized test scores.
  - Average precipitation levels.
  - Average price of certain stocks in the stock market.
  - The thickness of metal plates from a factory (can't be negative but still a good approximation since tail isn't close to 0).

Normal Distribution

- Examples include heights, weights, and other physical characteristics (the famous 1903 *Biometrika* article “On the Laws of Inheritance in Man” discussed many examples of this sort), measurement errors in scientific experiments, anthropometric measurements on fossils, reaction times in psychological experiments, measurements of intelligence and aptitude, scores on various tests, and numerous economic measures and indicators.
- The Central Limit Theorem in statistics indicates that random variables that are the sum of a number of component variables follow the normal distribution.

Normal Distribution

- The density function for a normally distributed random variable with mean μ and standard deviation σ, is

$$n(x;\mu,\sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty \le x \le \infty$$

Again *e* denotes the base of the natural logarithm system and equals approximately 2.71828, and π represents the familiar mathematical constant with approximate value 3.14159.

- As before, μ determines the center of the bell-shaped curve, while σ determines the spread. Notes:
  - The curve is symmetric about the mean.
  - As expected, the the total area under the curve equals 1. However, the function cannot be integrated in closed form.
  - The mode (where the curve is at a maximum) is at $x = \mu$.
  - The curve approaches the y axis as x moves away from μ.
  - The curve has a point of inflection at $x = \mu \pm \sigma$.

Normal Distribution Curve

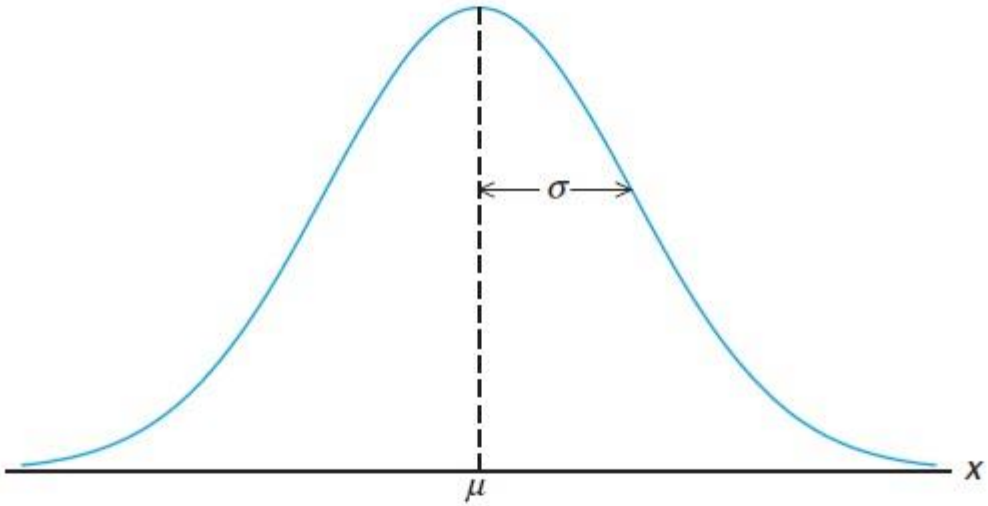


Figure 6.2: The normal curve.

The Normal Distribution

- The statement that X is normally distributed with parameters μ and σ^2 is often abbreviated $X \sim N(\mu, \sigma^2)$.

  Clearly $f(x; \mu, \sigma) \geq 0$, but a somewhat complicated calculus argument must be used to verify that $\int_{-\infty}^{\infty} f(x; \mu, \sigma)\, dx = 1$.

  It can be shown that $E(X) = \mu$ and $V(X) = \sigma^2$, so the parameters are the mean and the standard deviation of X.

The Normal Distribution

- Figure 4.13 presents graphs of $f(x; \mu, \sigma)$ for several different (μ, σ) pairs.

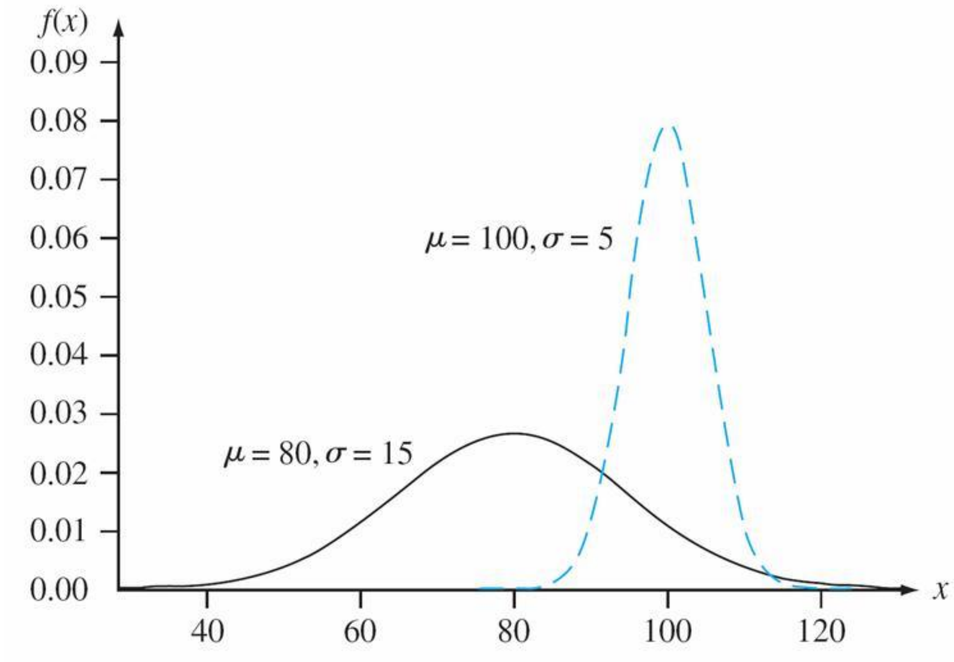
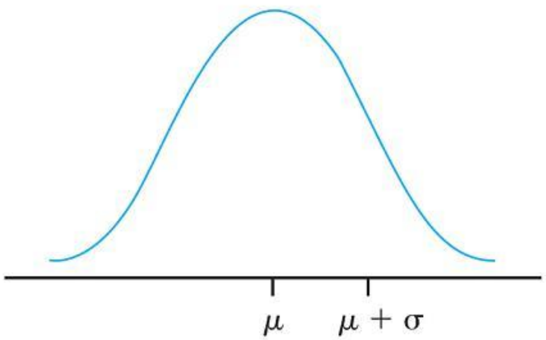


Figure 4.13(a)



Visualizing μ and σ for a normal distribution

Figure 4.13(b)

The Normal Distribution

- Each density curve is symmetric about μ and bell-shaped, so the center of the bell (point of symmetry) is both the mean of the distribution and the median.

  The mean μ is a *location parameter*, since changing its value rigidly shifts the density curve to one side or the

- other; σ is referred to as a *scale parameter*, because changing its value stretches or compresses the curve horizontally without changing the basic

The Normal Distribution

- The inflection points of a normal curve (points at which the curve changes from turning downward to turning upward) occur at $\mu - \sigma$ and $\mu + \sigma$. Thus the value of *s* can be visualized as the distance from the mean to these inflection points.
- A large value of *s* corresponds to a density curve that is quite spread out about μ, whereas a small value yields a highly concentrated curve
- The larger the value of σ, the more likely it is that a value of X far from the mean may be observed.

The Standard Normal Distribution

- The computation of $P(a \le X \le b)$ when X is a normal rv with parameters μ and σ requires evaluating

$$\int_a^b \frac{1}{\sqrt{2\pi}\sigma} e^{-(x-\mu)^2/(2\sigma^2)}\,dx \qquad (4.4)$$

None of the standard integration techniques can be used to accomplish this. Instead, for $\mu = 0$ and $\sigma = 1$, Expression (4.4) has been calculated using numerical techniques and tabulated for certain values of a and b.

This table can also be used to compute probabilities for any other values of μ and σ under consideration.

The Standard Normal Distribution

- **Definition**

The normal distribution with parameter values $\mu = 0$ and $\sigma = 1$ is called the **standard normal distribution.** A random variable having a standard normal distribution is called a **standard normal random variable** and will be denoted by Z. The pdf of Z is

$$f(z; 0, 1) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2} \quad -\infty < z < \infty$$

The graph of $f(z; 0, 1)$ is called the *standard normal* (or z) curve. Its inflection points are at 1 and -1. The cdf of Z is $P(Z \leq z) = \int_{-\infty}^{z} f(y; 0, 1)\, dy$, which we will denote by $\Phi(z)$.

Standard Normal Distribution

- Since it's not feasible to publish tables for every possible value of μ and σ, we define the standard normal distribution to be a normal distribution with $\mu = 0$ and $\sigma = 1$, and publish a table for that.
  - Then, every normal probability question can be reduced to an equivalent question about the standard normal distribution.
- A standard normal random variable can be from any normal random variable with a simple transformation:

$$Z = \frac{X - \mu}{\sigma}$$

Intuitively, Z shows the number of standard deviations above or below (if negative) the mean.

The Standard Normal Distribution

- The standard normal distribution almost never serves as a model for a naturally arising population.

  Instead, it is a reference distribution from which information
  about other normal distributions can be obtained.

- Appendix Table A.3 gives $\Phi(z) = P(Z \leq z)$, the area under the standard normal density curve to the left of z, for
  $z = -3.49, -3.48, \ldots, 3.48, 3.49$.

Normal Probability Table

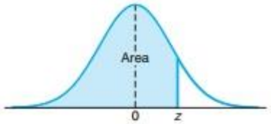
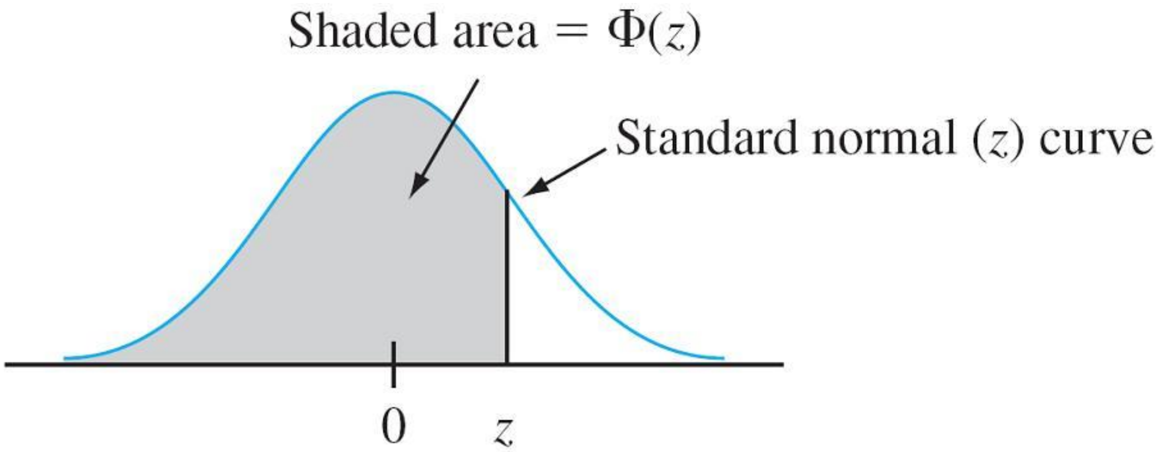


Table A.3 Areas under the Normal Curve

z	.00	.01	.02	.03	.04	.05	.06	.07	.08	.09
−3.4	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0002
−3.3	0.0005	0.0005	0.0005	0.0004	0.0004	0.0004	0.0004	0.0004	0.0004	0.0003
−3.2	0.0007	0.0007	0.0006	0.0006	0.0006	0.0006	0.0006	0.0005	0.0005	0.0005
−3.1	0.0010	0.0009	0.0009	0.0009	0.0008	0.0008	0.0008	0.0008	0.0007	0.0007
−3.0	0.0013	0.0013	0.0013	0.0012	0.0012	0.0011	0.0011	0.0011	0.0010	0.0010
−2.9	0.0019	0.0018	0.0018	0.0017	0.0016	0.0016	0.0015	0.0015	0.0014	0.0014
−2.8	0.0026	0.0025	0.0024	0.0023	0.0023	0.0022	0.0021	0.0021	0.0020	0.0019
−2.7	0.0035	0.0034	0.0033	0.0032	0.0031	0.0030	0.0029	0.0028	0.0027	0.0026
−2.6	0.0047	0.0045	0.0044	0.0043	0.0041	0.0040	0.0039	0.0038	0.0037	0.0036
−2.5	0.0062	0.0060	0.0059	0.0057	0.0055	0.0054	0.0052	0.0051	0.0049	0.0048
−2.4	0.0082	0.0080	0.0078	0.0075	0.0073	0.0071	0.0069	0.0068	0.0066	0.0064
−2.3	0.0107	0.0104	0.0102	0.0099	0.0096	0.0094	0.0091	0.0089	0.0087	0.0084
−2.2	0.0139	0.0136	0.0132	0.0129	0.0125	0.0122	0.0119	0.0116	0.0113	0.0110
−2.1	0.0179	0.0174	0.0170	0.0166	0.0162	0.0158	0.0154	0.0150	0.0146	0.0143
−2.0	0.0228	0.0222	0.0217	0.0212	0.0207	0.0202	0.0197	0.0192	0.0188	0.0183
−1.9	0.0287	0.0281	0.0274	0.0268	0.0262	0.0256	0.0250	0.0244	0.0239	0.0233
−1.8	0.0359	0.0351	0.0344	0.0336	0.0329	0.0322	0.0314	0.0307	0.0301	0.0294
−1.7	0.0446	0.0436	0.0427	0.0418	0.0409	0.0401	0.0392	0.0384	0.0375	0.0367
−1.6	0.0548	0.0537	0.0526	0.0516	0.0505	0.0495	0.0485	0.0475	0.0465	0.0455
−1.5	0.0668	0.0655	0.0643	0.0630	0.0618	0.0606	0.0594	0.0582	0.0571	0.0559
−1.4	0.0808	0.0793	0.0778	0.0764	0.0749	0.0735	0.0721	0.0708	0.0694	0.0681
−1.3	0.0968	0.0951	0.0934	0.0918	0.0901	0.0885	0.0869	0.0853	0.0838	0.0823

The Standard Normal Distribution

- Figure 4.14 illustrates the type of cumulative area (probability) tabulated in Table A.3. From this table, various other probabilities involving *Z* can be calculated.



Standard normal cumulative areas tabulated in Appendix Table A.3

Figure 4.14

Example 4.13

- Let's determine the following standard normal probabilities:

(a) $P(Z \leq 1.25)$, (b) $P(Z > 1.25)$,

(c) $P(Z \leq -1.25)$, and (d) $P(-.38 \leq Z \leq 1.25)$.

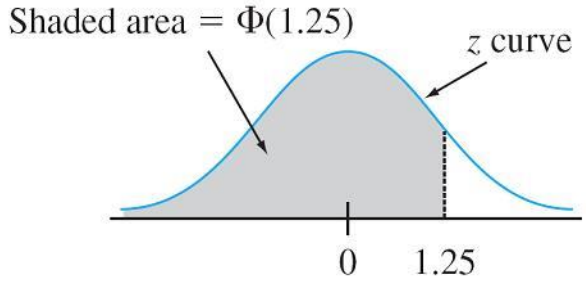
a. $P(Z \leq 1.25) = \Phi(1.25)$, a probability that is tabulated in Appendix Table A.3 at the intersection of the row marked 1.2 and the column marked .05.

The number there is .8944, so $P(Z \leq 1.25) = .8944$.

Example 4.13

cont'd

Figure 4.15(a) illustrates this probability.



Normal curve areas (probabilities) for Example 13

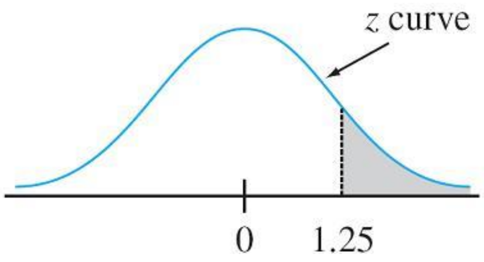
Figure 4.15(a)

b. $P(Z > 1.25) = 1 - P(Z \leq 1.25) = 1 - \Phi(1.25)$, the area under the z curve to the right of 1.25 (an upper-tail area). Then $\Phi(1.25) = .8944$ implies that $P(Z > 1.25) = .1056$.

cont'd

Example 4.13

Since Z is a continuous rv, $P(Z \geq 1.25) = .1056$. See Figure 4.15(b).



Normal curve areas (probabilities) for Example 13

Figure 4.15(b)

c. $P(Z \leq -1.25) = \Phi(-1.25)$, a lower-tail area. Directly from Appendix Table A.3, $\Phi(-1.25) = .1056$. By symmetry of the z curve, this is the same answer as in part (b).

Example 4.13

cont'd

- **d.** $P(-.38 \le Z \le 1.25)$ is the area under the standard normal curve above the interval whose left endpoint is –.38 and whose right endpoint is 1.25.

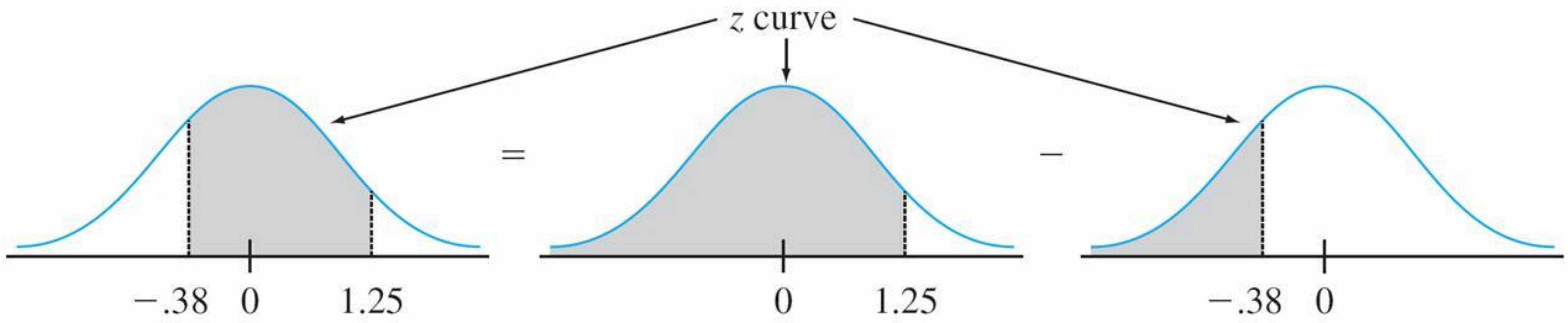
- From Section 4.2, if X is a continuous rv with cdf $F(x)$, then $P(a \le X \le b) = F(b) - F(a)$.
- Thus $P(-.38 \le Z \le 1.25) = \Phi(1.25) - \Phi(-.38)$

$$= .8944 - .3520$$

$$= .5424$$

Example 4.13

cont'd



$P(-.38 \leq Z \leq 1.25)$ as the difference between two cumulative areas

Figure 4.16

Using the Normal Distribution Table (A.3)

- The table shows F(z), which is?
  - The cumulative standard normal distribution. Which means?
- We use the key fact that the normal distribution is symmetric around to mean, and that the area under the curve equals 1, to find any desired probability.
- To use the table, first convert the normal random variable to a standard normal random variable.
  - Ex: If a normal random variable X has $\mu = 50$ and $\sigma = 25$, find P(X > 12).
  - Converting to standard normal, we have P(X > 12) = P(Z > -1.52). (12 is 1.52 standard deviations below the mean.)
  - By table A.3, F(-1.52) = .0643. So P(Z > -1.52) = 1 - .0643 = .9357.
  - Or, by symmetry, P(Z > -1.52) = P(Z < +1.52) = .9357.
  - So, P(X > 12) = .9357.

More Normal Distribution Calculations

- Use properties of a cumulative distribution to calculate the probability of being within a given range.
  - Ex: If a normal random variable X has $\mu = 50$ and $\sigma = 25$, find $P(62 > X > 32)$. First step?
  - Convert to standard normal.
  - $P(62 > X > 32) = P(0.48 > Z > -0.72)$. Next?
  - $P(0.48 > Z > -0.72) = P(Z \leq 0.48) - P(Z \leq -0.72) = ?$
  - $P(Z \leq 0.48) - P(Z \leq -0.72) = 0.6844 - 0.2358 = .4486$.
- Reverse calculation: starting with a probability, find X.
  - Using μ and σ above, what values of X centered above and below the mean have p = .80 of falling between them. First?
  - Find Z values with 80% of the area between them. $z = \pm 1.28$.
  - Next convert from standard normal back to X.
  - So, the limits are $x = 50 \pm 1.28(25) = (18, 82)$.

Example 4.13

- **e.** $P(Z \leq 5) = \Phi(5)$, the cumulative area under the z curve to the left of 5. This probability does not appear in the table because the last row is labeled 3.4. However, the last entry in that row is $\Phi(3.49) = .9998$. That is, essentially all of the area under the curve lies to the left of 3.49 (at most 3.49 standard deviations to the right of the mean). Therefore we conclude that $P(Z \leq 5) \approx 1$.

Percentiles of the Standard Normal Distribution

- For any p between 0 and 1, Appendix Table A.3 can be used to obtain the $(100p)$th percentile of the standard normal distribution.

Example 4.14

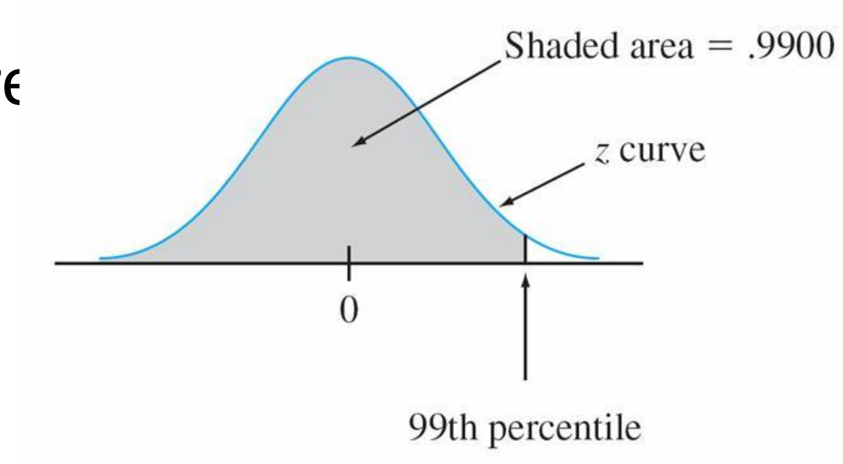
- The 99th percentile of the standard normal distribution is that value on the horizontal axis such that the area under the *z* curve to the left of the value is .9900.
- Appendix Table A.3 gives for fixed *z* the area under the standard normal curve to the left of *z*, whereas here we have the area and want the value of *z*. This is the "inverse" problem to $P(Z \leq z) = ?$
- so the table is used in an inverse fashion: Find in the middle of the table .9900; the row and column in which it lies identify the 99th *z* percentile.

Example 4.14

cont'd

- Here .9901 lies at the intersection of the row marked 2.3 and column marked .03, so the 99th percentile is (approximately) $z = 2.33$.

- (See Figure



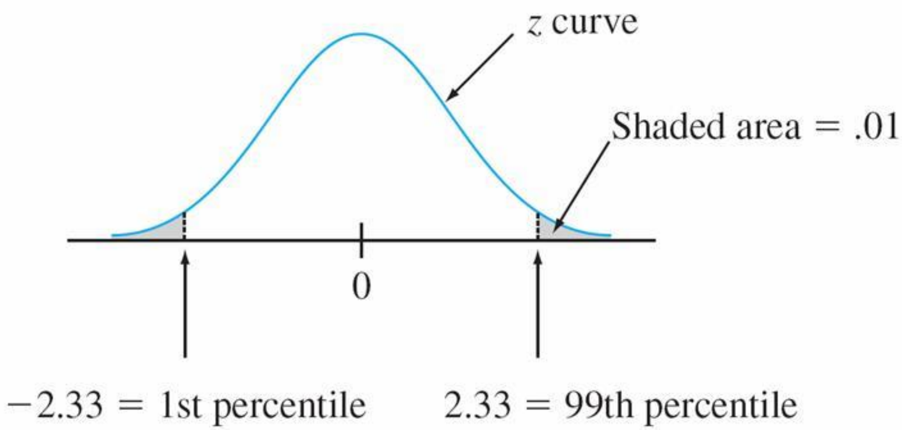
Finding the 99th percentile

Figure 4.17

Example 4.14

cont'd

- By symmetry, the first percentile is as far below 0 as the 99th is above 0, so equals –2.33 (1% lies below the first and also above the 99th). (See Figure 4.18.)



The relationship between the 1st and 99th percentiles

Figure 4.18

Percentiles of the Standard Normal Distribution

- In general, the (100p)th percentile is identified by the row and column of Appendix Table A.3 in which the entry p is found (e.g., the 67th percentile is obtained by finding .6700 in the body of the table, which gives $z = .44$).

  If p does not appear, the number closest to it is often used, although linear interpolation gives a more accurate answer.

Percentiles of the Standard Normal Distribution

- For example, to find the 95th percentile, we look for .9500 inside the table.

  Although .9500 does not appear, both .9495 and .9505 do, corresponding to z = 1.64 and 1.65, respectively.

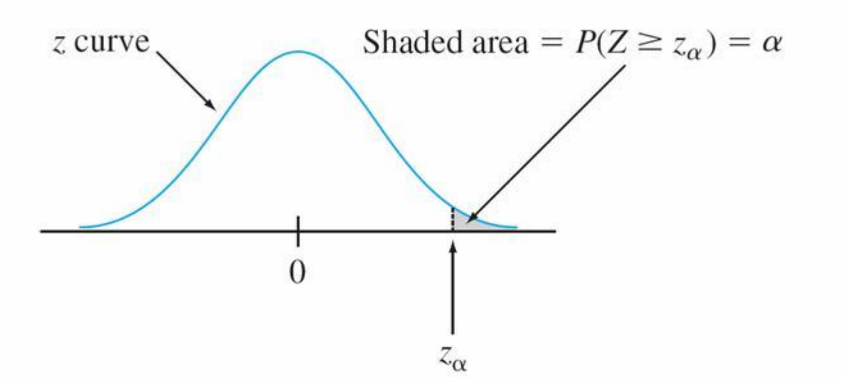
- Since .9500 is halfway between the two probabilities that do appear, we will use 1.645 as the 95th percentile and –1.645 as the 5th percentile.

z_α Notation for z Critical Values

- In statistical inference, we will need the values on the horizontal z axis that capture certain small tail

Notation

z_α will denote the value on the z axis for which α of the area under the z curve lies to the right of z_α. (See Figure 4.19.)



z_α notation Illustrated

Figure 4.19

z_α Notation for z Critical Values

- For example, $z_{.10}$ captures upper-tail area .10, and $z_{.01}$ captures upper-tail area .01.

  Since α of the area under the z curve lies to the right of z_α, $1 - \alpha$ of the area lies to its left. Thus z_α *is the* $100(1 - \alpha)$*th percentile of the standard normal distribution*.

  By symmetry the area under the standard normal curve to the left of $-z_\alpha$ is also α. The z_α's are usually referred to as ***z* critical values.**

z_α Notation for z Critical Values

- Table 4.1 lists the most useful z percentiles and z_α values.

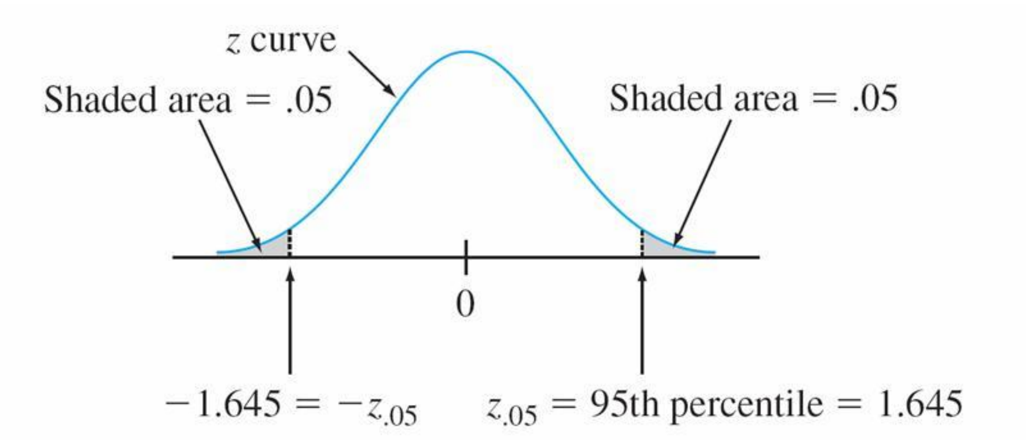
Percentile	90	95	97.5	99	99.5	99.9	99.95
α (tail area)	.1	.05	.025	.01	.005	.001	.0005
$z_\alpha = 100(1-\alpha)$th percentile	1.28	1.645	1.96	2.33	2.58	3.08	3.27

Standard Normal Percentiles and Critical Values

Table 4.1

Example 4.15

- $z_{.05}$ is the 100(1 – .05)th = 95th percentile of the standard normal distribution, so $z_{.05}$ = 1.645.
- The area under the standard normal curve to the left of $-z_{.05}$ is al



Finding $z_{.05}$

Figure 4.20

Nonstandard Normal Distributions

- When $X \sim N(\mu, \sigma^2)$, probabilities involving X are computed by “standardizing.” The **standardized variable** is $(X - \mu)/\sigma$.

- Subtracting μ shifts the mean from μ to zero, and then dividing by σ scales the variable so that the standard deviation is 1 rather than σ.

Nonstandard Normal Distributions

- **Proposition**

If X has a normal distribution with mean μ and standard deviation σ, then

$$Z = \frac{X - \mu}{\sigma}$$

has a standard normal distribution. Thus

$$P(a \le X \le b) = P\left(\frac{a-\mu}{\sigma} \le Z \le \frac{b-\mu}{\sigma}\right)$$

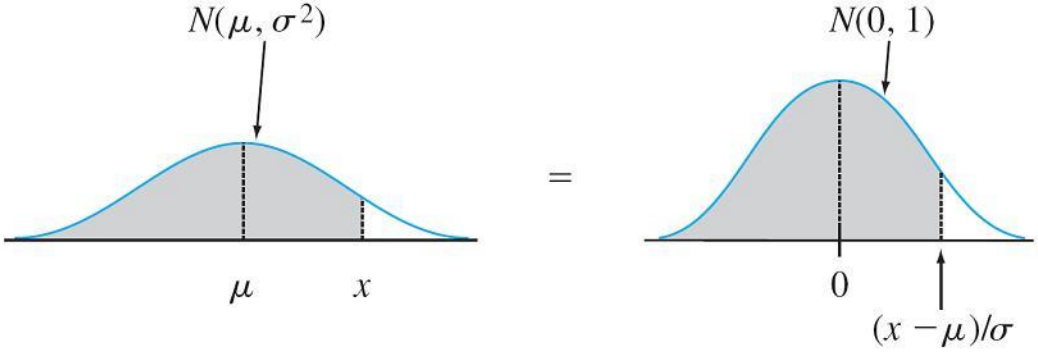
$$= \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

$$P(X \le a) = \Phi\left(\frac{a-\mu}{\sigma}\right) \qquad P(X \ge b) = 1 - \Phi\left(\frac{b-\mu}{\sigma}\right)$$

Nonstandard Normal Distributions

- The key idea of the proposition is that by standardizing, any
  probability involving *X* can be expressed as a probability involving a standard normal rv *Z*, so that Appendix Table A.3 can be used.

  This is



Equality of nonstandard and standard normal curve areas

Figure 4.21

Nonstandard Normal Distributions

- The proposition can be proved by writing the cdf of $Z = (X - \mu)/\sigma$ as

$$P(Z \leq z) = P(X \leq \sigma z + \mu) = \int_{-\infty}^{\sigma z + \mu} f(x; \mu, \sigma)dx$$

Using a result from calculus, this integral can be differentiated with respect to *z* to yield the desired pdf $f(z; 0, 1)$.

Example 4.16

- The time that it takes a driver to react to the brake lights on a decelerating vehicle is critical in helping to avoid rear-end collisions.
- The article “Fast-Rise Brake Lamp as a Collision-Prevention Device” (*Ergonomics*, 1993: 391–395) suggests that reaction time for an in-traffic response to a brake signal from standard brake lights can be modeled with a normal distribution having mean value 1.25 sec and standard deviation of .46 sec.

Example 4.16

cont'd

- What is the probability that reaction time is between 1.00 sec and 1.75 sec? If we let *X* denote reaction time, then standardizing gives

$$1.00 \leq X \leq 1.75$$

- if and only if

$$\frac{1.00 - 1.25}{.46} \leq \frac{X - 1.25}{.46} \leq \frac{1.75 - 1.25}{.46}$$

Thus

$$P(1.00 \leq X \leq 1.75) = P\left(\frac{1.00 - 1.25}{.46} \leq Z \leq \frac{1.75 - 1.25}{.46}\right)$$

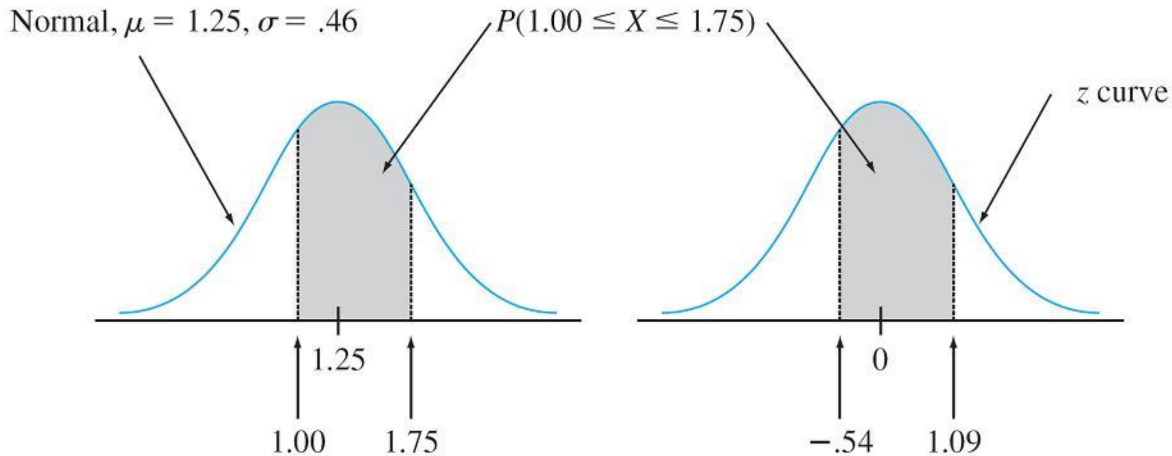
Example 4.16

cont'd

$$= P(-.54 \leq Z \leq 1.09) = \Phi(1.09) - \Phi(-.54)$$

$$= .8621 - .2946 = .5675$$

- This is illustrated in Figure 4.22



Normal curves for Example 16

Figure 4.22

Example 4.16

- Similarly, if we view 2 sec as a critically long reaction time, the probability that actual reaction time will exceed this value is

$$P(X > 2) = P\left(Z > \frac{2 - 1.25}{.46}\right) = P(Z > 1.63) = 1 - \Phi(1.63) = .0516$$

Nonstandard Normal Distributions

- These results are often reported in percentage form and referred to as the *empirical rule* (because empirical evidence has shown that histograms of real data can very frequently be approximated by normal curves).

If the population distribution of a variable is (approximately) normal, then

1. Roughly 68% of the values are within 1 SD of the mean.
2. Roughly 95% of the values are within 2 SDs of the mean.
3. Roughly 99.7% of the values are within 3 SDs of the mean.

- It is indeed unusual to observe a value from a normal population that is much farther than 2 standard deviations from *m*. These results will be important in the development of hypothesis-testing procedures in later chapters

Percentiles of an Arbitrary Normal Distribution

- The (100*p*)th percentile of a normal distribution with mean μ and standard deviation σ is easily related to the (100*p*)th percentile of the standard normal distribution.

Proposition

$$\begin{array}{c}(100p)\text{th percentile}\\ \text{for normal } (\mu, \sigma)\end{array} = \mu + \left[\begin{array}{c}(100p)\text{th for}\\ \text{standard normal}\end{array}\right] \cdot \sigma$$

- Another way of saying this is that if *z* is the desired percentile for the standard normal distribution, then the desired percentile for the normal (μ, σ) distribution is *z* standard deviations from μ.

Example 4.18

- The authors of "Assessment of Lifetime of Railway Axle" (*Intl. J. of Fatigue*, (2013: 40–46) used data collected from an experiment with a specified initial crack length and number of loading cycles to propose a normal distribution with mean value 5.496 mm and standard deviation .067 mm for the rv X = final crack depth.

- For this model, what value of final crack depth would be exceeded by only .5% of all cracks under these circumstances? Let c denote the requested value. Then the desired condition is that $P(X > c) = .005$, or, equivalently, that $P(X \leq c) = .995$.

Example 4.18

cont'd

- Thus *c* is the 99.5th percentile of the normal distribution with $\mu = 5.496$ and $\sigma = .067$. The 99.5th percentile of the standard normal distribution is 2.58, so

$$c = \eta(.995) = 5.496 + (2.58)(.067) = 5.496 + .173 = 5.669 \text{ mm}$$

This is illustrated in Figure 4.23.

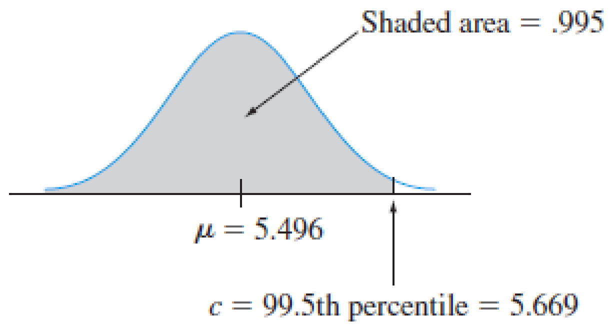


Figure 4.23 Distribution of final crack depth for Example 4.18

The Normal Distribution and Discrete Populations

- The normal distribution is often used as an approximation to the distribution of values in a discrete population.

- In such situations, extra care should be taken to ensure that probabilities are computed in an accurate manner.

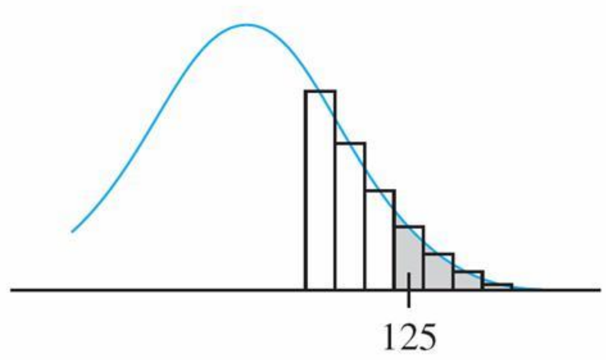
Example 4.19

- IQ in a particular population (as measured by a standard test) is known to be approximately normally distributed with $\mu = 100$ and $\sigma = 15$.
- What is the probability that a randomly selected individual has an IQ of at least 125?
- Letting X = the IQ of a randomly chosen person, we wish $P(X \geq 125)$. The temptation here is to standardize $X \geq 125$ as in previous examples. However, the IQ population distribution is actually discrete, since IQs are integer-valued.

Example 4.19

cont'd

- So the normal curve is an approximation to a discrete probability histogram, as pictured in Figure 4.24.



A normal approximation to a discrete distribution

Figure 4.24

- The rectangles of the histogram are *centered* at integers, so IQs of at least 125 correspond to rectangles beginning at 124.5, as shaded in Figure 4.24.

Example 4.19

cont'd

- Thus we really want the area under the approximating normal curve to the right of 124.5.

- Standardizing this value gives $P(Z \geq 1.63) = .0516$, whereas standardizing 125 results in $P(Z \geq 1.67) = .0475$.

- The difference is not great, but the answer .0516 is more accurate. Similarly, $P(X = 125)$ would be approximated by the area between 124.5 and 125.5, since the area under the normal curve above the single value 125 is zero.

Example 4.19

cont'd

- The correction for discreteness of the underlying distribution in Example 19 is often called a **continuity correction.**

- It is useful in the following application of the normal distribution to the computation of binomial probabilities.

Normal Approximation to Binomial

- If X binomial with parameters n and p, then

$$P(X \leq x) \;\approx\; P\left(Z \leq \frac{x + 0.5 - np}{\sqrt{npq}} \right)$$

  for n $\to \infty$. This is a good approximation:
  - If n is large and p is not too close to 0 or 1.
  - Or, if p is reasonably close to 1/2, even when n is small.
  - To find the probability that x = k?
  - Use the normal distribution for the area x = k $\pm$ (1/2).
  - For k = 0, we use P(x $\leq$ +0.5) so probability will sum to 1.
- Approximation works well if both np and nq $\geq$ 5.
  - The larger n is, the better the approximation will be.

Approximating the Binomial Distribution

- Figure 4.25 displays a binomial probability histogram for the binomial distribution with n = 25, p = .6, for which $\mu = 25(.6) = 15$ and $\sigma = \sqrt{25(.6)(.4)} = 2.449$

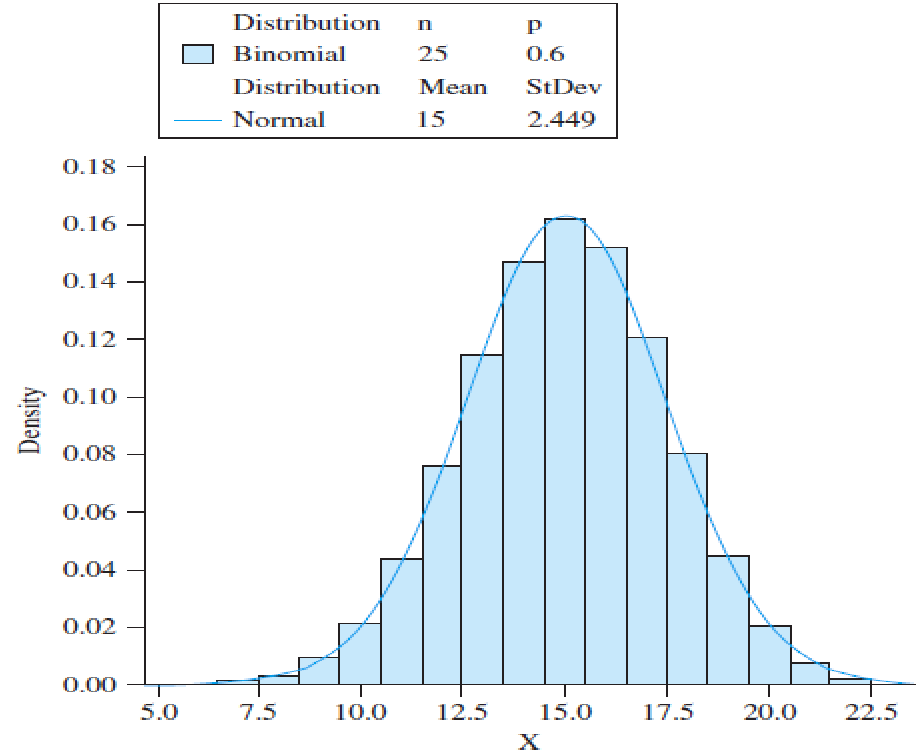


Figure 4.25 Binomial probability histogram for $n = 25$, $p = .6$ with normal approximation curve superimposed

Approximating the Binomial Distribution

- A normal curve with this μ and σ has been superimposed on the probability histogram.

  Although the probability histogram is a bit skewed (because $p \neq .5$), the normal curve gives a very good approximation, especially in the middle part of the picture.

  The area of any rectangle (probability of any particular X value) except those in the extreme tails can be accurately approximated by the corresponding normal curve area.

Approximating the Binomial Distribution

- For example,

$P(X = 10) = B(10; 25, .6) - B(9; 25, .6) = .021$,

whereas the area under the normal curve between 9.5 and 10.5 is $P(-2.25 \leq Z \leq -1.84) = .0207$.

More generally, as long as the binomial probability histogram is not too skewed, binomial probabilities can be well approximated by normal curve areas.

It is then customary to say that X has approximately a normal distribution.

Approximating the Binomial Distribution

Let X be a binomial rv based on n trials with success probability p. Then if the binomial probability histogram is not too skewed, X has approximately a normal distribution with $\mu = np$ and $\sigma = \sqrt{npq}$. In particular, for $x =$ a possible value of X,

$$P(X \le x) = B(x, n, p) \approx \begin{pmatrix} \text{area under the normal curve} \\ \text{to the left of } x + .5 \end{pmatrix}$$
$$= \Phi\left(\frac{x + .5 - np}{\sqrt{npq}}\right)$$

In practice, the approximation is adequate provided that both $np \ge 10$ and $nq \ge 10$ (i.e., the expected number of successes and the expected number of failures are both at least 10), since there is then enough symmetry in the underlying binomial distribution.

Approximating the Binomial Distribution

- A direct proof of this result is quite difficult. In the next chapter we'll see that it is a consequence of a more general result called the Central Limit Theorem.

  In all honesty, this approximation is not so important for probability calculation as it once was.

- This is because software can now calculate binomial probabilities exactly for quite large values of n.

Example 4.20

- Suppose that 25% of all students at a large public university receive financial aid.
- Let X be the number of students in a random sample of size 50 who receive financial aid, so that p = .25.
  Then μ = 12.5 and σ = 3.06.
- Since np = 50(.25) = 12.5 $\geq$ 10 and np = 37.5 $\geq$ 10, the approximation can safely be applied.

Example 4.20

cont'd

- The probability that at most 10 students receive aid is

$$P(X \leq 10) = B(10; 50, .25) \approx \Phi\left(\frac{10 + .5 - 12.5}{3.06}\right)$$

$$= \Phi(-.65) = .2578$$

- Similarly, the probability that between 5 and 15 (inclusive) of the selected students receive aid is

- $P(5 \leq X \leq$ $\approx \Phi\left(\frac{15.5 - 12.5}{3.06}\right) - \Phi\left(\frac{4.5 - 12.5}{3.06}\right) = .8320$

Example 4.20

cont'd

- The exact probabilities are .2622 and .8348, respectively, so the approximations are quite good.

- In the last calculation, the probability $P(5 \leq X \leq 15)$ is being approximated by the area under the normal curve between 4.5 and 15.5—the continuity correction is used for both the upper and lower limits.

Distribution Approximation Notes

- Continuity correction: Suppose a population of resistors follows a normal distribution, but that resistance measurements are in integer values of ohms.
  - Here we assume that any resistance from 42.5 to 43.5 = 43.
  - So the probability of any outcome can be calculated using the normal distribution.
  - This example uses a continuity correction to approximate a discrete distribution with a continuous distribution.
- If n is large and p is very small (or very large), the binomial distribution can be approximated by what?
  - The Poisson distribution.
- What is μ and σ for the binomial distribution?
  - $\mu = np$.
  - $\sigma^2 = npq$.

Assessing Normality

- Not all continuous random variables are normally distributed
- It is important to evaluate how well the data set is approximated by a normal distribution

Assessing Normality

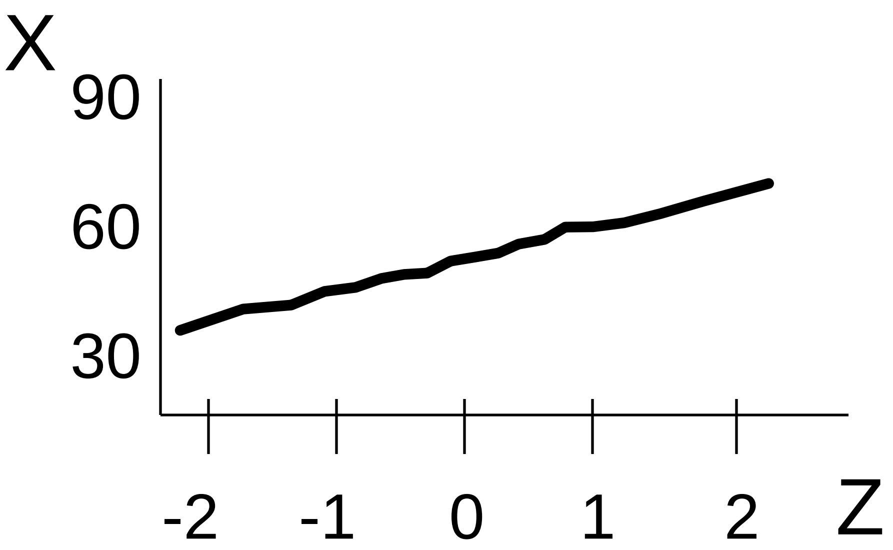
(continued)

- Construct charts or graphs
  - For small- or moderate-sized data sets, do a stem-and-leaf display and box-and-whisker plot look symmetric?
  - For large data sets, does the histogram or polygon appear bell-shaped?
- Compute descriptive summary measures
  - Do the mean, median and mode have similar values?
  - Is the interquartile range approximately 1.33 σ?
  - Is the range approximately 6 σ?
- Evaluate normal probability plot

The Normal Probability Plot

(continued)

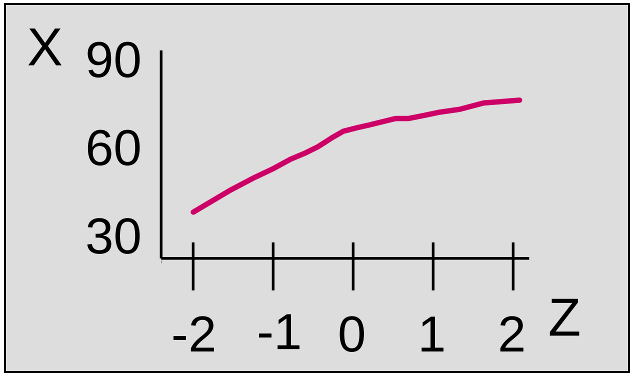
A normal probability plot for data from a normal distribution will be approximately linear:



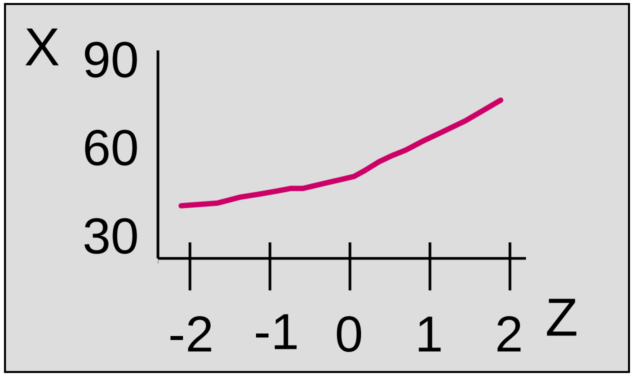
Normal Probability Plot for Non-Normal Distributions

(continued)

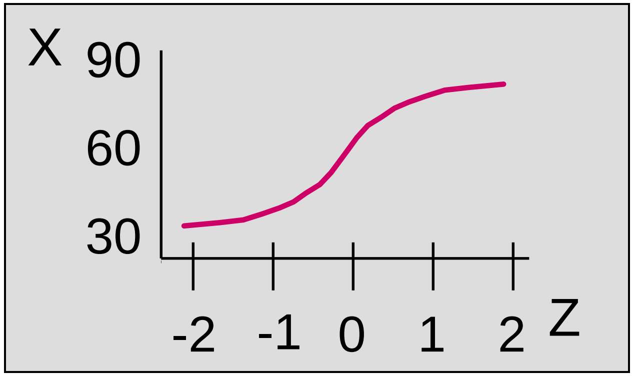
Left-Skewed



Right-Skewed



Rectangular



Nonlinear plots indicate a deviation from normality

Exponential Distribution

- An exponential distribution with parameter λ, is defined as

$$f(t) = \lambda e^{-\lambda t} \qquad for\ t > 0$$
$$= 0 \qquad elsewhere$$

then $$\mu = \frac{1}{\lambda}$$

and $$\sigma^2 = \frac{1}{\lambda^2}$$

Exponential distribution notes:

- The mean equals the standard deviation.
- Note that $\lambda = (1/\beta)$, where β is the parameter used in the book for the exponential distribution.

The Exponential Distributions

- The family of exponential distributions provides probability models that are very widely used in engineering and science disciplines.

X is said to have an **exponential distribution** with (scale) parameter λ ($\lambda > 0$) if the pdf of X is

$$f(x; \lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.5}$$

The Exponential Distributions

- Some sources write the exponential pdf in the form $(1/\beta)e^{-x/\beta}$, so that $\beta = 1/\lambda$. The expected value of an exponentially distributed random variable *X* is

$$E(X) = \int_0^{\infty} x\lambda e^{-\lambda x}\, dx$$

- Obtaining this expected value necessitates doing an integration by parts. The variance of *X* can be computed using the fact that $V(X) = E(X^2) - [E(X)]^2$.

- The determination of $E(X^2)$ requires integrating by parts twice in succession.

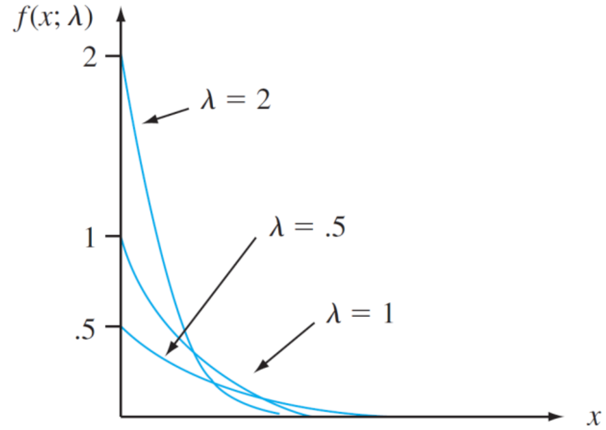
The Exponential Distributions

- The results of these integrations are as follows:

$$\mu = \frac{1}{\lambda} \qquad \sigma^2 = \frac{1}{\lambda^2}$$

- Both the mean and standard deviation of the exponential distribution equal

- Graphs of several exponential pdf's are illustrated in Figure 4



Exponential density curves

Figure 4.26

The Exponential Distributions

- The exponential pdf is easily integrated to obtain the cdf.

$$F(x; \lambda) = \begin{cases} 0 & x < 0 \\ 1 - e^{-\lambda x} & x \geq 0 \end{cases}$$

Example 4.21

- The article "Probabilistic Fatigue Evaluation of Riveted Railway Bridges" (*J. of Bridge Engr.,* 2008: 237–244) suggested the exponential distribution with mean value 6 MPa as a model for the distribution of stress range in certain bridge connections.

- Let's assume that this is in fact the true model. Then
  $E(X) = 1/\lambda = 6$ implies that $\lambda = .1667$.

Example 4.21

cont'd

The probability that stress range is at most 10 MPa is

$$P(X \leq 10) = F(10\ ;\ .1667)$$

$$= 1 - e^{-(.1667)(10)}$$

$$= 1 - .189$$

$$= .811$$

Example 4.21

cont'd

The probability that stress range is between 5 and 10 MPa is

$$P(5 \leq X \leq 10) = F(10; .1667) - F(5; .1667)$$

$$= (1 - e^{-1.667}) - (1 - e^{-.8335})$$

$$= .246$$

The Exponential Distributions

- The exponential distribution is frequently used as a model for the distribution of times between the occurrence of successive events, such as customers arriving at a service facility or calls coming in to a switchboard.

- The reason for this is that the exponential distribution is closely related to the Poisson process discussed in Chapter 3.

The Exponential Distributions

- **Proposition**

Suppose that the number of events occurring in any time interval of length t has a Poisson distribution with parameter αt (where α, the rate of the event process, is the expected number of events occurring in 1 unit of time) and that numbers of occurrences in nonoverlapping intervals are independent of one another. Then the distribution of elapsed time between the occurrence of two successive events is exponential with parameter $\lambda = \alpha$.

The Exponential Distributions

Although a complete proof is beyond the scope of the text, the result is easily verified for the time X_1 until the first event occurs:

$$P(X_1 \leq t) = 1 - P(X_1 > t) = 1 - P\,[\text{no events in } (0, t)]$$

$$= 1 - \frac{e^{-\alpha t} \cdot (\alpha t)^0}{0!} = 1 - e^{-\alpha t}$$

which is exactly the cdf of the exponential distribution.

Example 4.22

Suppose that calls to a rape crisis center in a certain county occur according to a Poisson process with rate α = .5 call per day.

Then the number of days X between successive calls has an exponential distribution with parameter value .5, so the probability that more than 2 days elapse between calls is

$$P(X > 2) = 1 - P(X \leq 2)$$

$$= 1 - F(2; .5)$$

$$= e^{-(.5)(2)}$$

$$= .368$$

The expected time between successive calls is 1/.5 = 2 days.

Exponential and Poisson Distributions

- For a Poisson process, if λ is the mean event arrival rate, what is the distribution of the number of events occurring in a time interval of length t?
  - It is a Poisson distribution with mean λt.
- It can also be shown for a Poisson process that the distribution of the time until the next event occurs is exponential with parameter λ.
- Similarly, an Erlang distribution with parameter α is the distribution of the time until α events occur.
  - Exponential distribution is a gamma distribution with $\alpha = 1$.
- There are many important applications of the exponential distribution (and the gamma distribution) in the areas of reliability and queuing theory.

The Exponential Distributions

- Another important application of the exponential distribution is to model the distribution of component lifetime.

- A partial reason for the popularity of such applications is the **“memoryless” property** of the exponential distribution.

- Suppose component lifetime is exponentially distributed with parameter λ.

The Exponential Distributions

- After putting the component into service, we leave for a period of t_0 hours and then return to find the component still working; what now is the probability that it lasts at least an additional t hours?
- In symbols, we wish $P(X \geq t + t_0 \mid X \geq t_0)$.

$$P(X \geq t + t_0 | X \geq t_0) = \frac{P[(X \geq t + t_0) \cap (X \geq t_0)]}{P(X \geq t_0)}$$

- By the definition of conditional probability,

The Exponential Distributions

- But the event $X \geq t_0$ in the numerator is redundant, since both events can occur if $X \geq t + t_0$ and only if. Therefore,

$$P(X \geq t + t_0 \mid X \geq t_0) = \frac{P(X \geq t + t_0)}{P(X \geq t_0)} = \frac{1 - F(t + t_0; \lambda)}{1 - F(t_0; \lambda)} = e^{-\lambda t}$$

- This conditional probability is identical to the original probability $P(X \geq t)$ that the component lasted t hours.

Exponential Dist'n - Memoryless Property

- What was the memoryless property for a Poisson process?
  - The number of occurrences in one time interval is independent of the number in any other disjoint time interval.
  - Thus future arrivals are completely independent of the past history of arrivals.
- How does the memoryless concept apply to the exponential distribution of the time between arrivals?
  - The memoryless property of the exponential distribution implies that

$$P(X \geq t_0 + t \mid X \geq t_0) = P(X \geq t)$$

Or, the probability of waiting longer than t time units for the next arrival is independent of how long we've already waited.

Exponential Distribution Example 1

Example: Customers arrive at the service counter at the rate of 15 per hour. What is the probability that the arrival time between consecutive customers is less than three minutes?

- The mean number of arrivals per hour is 15, so $\lambda = 15$
- Three minutes is .05 hours
- P(arrival time < .05) $= 1 - e^{-\lambda X} = 1 - e^{-(15)(.05)} = \boxed{.5276}$
- So there is a 52.76% probability that the arrival time between successive customers is less than three minutes

Exponential Distribution
Example 2

Suppose that a study of certain computer system reveals that the response time, in seconds, has an exponential distribution with a mean of 3 seconds.

a. What is the probability that response time exceeds 5 seconds? (answer = 0.1889)

b. What is the probability that response time exceeds 10 seconds? (answer = 0.0357)

Exponential Distribution
Example 3

The time (in hours) required to repair a machine is an exponentially distributed random variable with parameter $\lambda = 0.5$

a. What is the probability that repair time exceeds 2 hours?

b. What is the conditional probability that a repair takes at least 10 hours, given that its duration exceeds 9 hours?

Answers

- $P(X > 2) = 0.3679$
- $P(X \geq t_0+t \mid X \geq t_0) = P(X \geq t) = P(X \geq 1)$ 0.6065

The Exponential Distributions

- Thus *the distribution of additional lifetime is exactly the same as the original distribution of lifetime,* so at each point in time the component shows no effect of wear.

- In other words, the distribution of remaining lifetime is independent of current age.

Cumulative Exponential Distributions

- The cumulative distribution for the exponential is very simple. It is

$$F(t) = 1 - e^{-\lambda t} \qquad \textit{for } t > 0$$
$$= \ 0 \qquad \textit{elsewhere}$$

    - It is easiest to use the cumulative distribution to calculate exponential distribution probabilities.

The Gamma Function

The Gamma Function

- To define the family of gamma distributions, we first need to introduce a function that plays an important role in many branches of mathematics.

- **Definition**

For $\alpha > 0$, the **gamma function** $\Gamma(\alpha)$ is defined by

$$\Gamma(\alpha) = \int_0^\infty x^{\alpha-1}e^{-x}\,dx \tag{4.6}$$

The Gamma Function

- The most important properties of the gamma function are the following:

- **1.** For any $\alpha > 1$, $\Gamma(\alpha) = (\alpha - 1) \cdot \Gamma(\alpha)$ $(\alpha - 1)$ [via integration by parts]

- **2.** For any positive integer, *n*, $\Gamma(n) = (n - 1)!$

- **3.** $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

The Gamma Function

Now let

$$f(x;\alpha) = \begin{cases} \dfrac{x^{\alpha-1}e^{-x}}{\Gamma(\alpha)} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \qquad \textbf{(4.7)}$$

Then f(x; α) ≥ 0 Expression (4.6) implies that ,

$$\int_0^{\infty} f(x;\alpha)\,dx = \Gamma(\alpha)/\Gamma(\alpha) = 1$$

Thus *f*(*x*; a) satisfies the two basic properties of a pdf.

Gamma Distribution

- The gamma distribution with parameters α and β, is defined as

$$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}} \qquad for\ \ x > 0$$

$$= \quad 0 \qquad\qquad elsewhere$$

where $\alpha > 0$, $\beta > 0$, and $\Gamma(\alpha) = \int_0^{\infty} x^{\alpha-1} e^{-x} dx$

- Note that if α is an integer, $\Gamma(\alpha)$ simplifies to $(\alpha-1)$!
  then
  and

$$\mu = \alpha\beta$$

$$\sigma^2 = \alpha\beta^2$$

The Gamma Distribution

- **Definition**

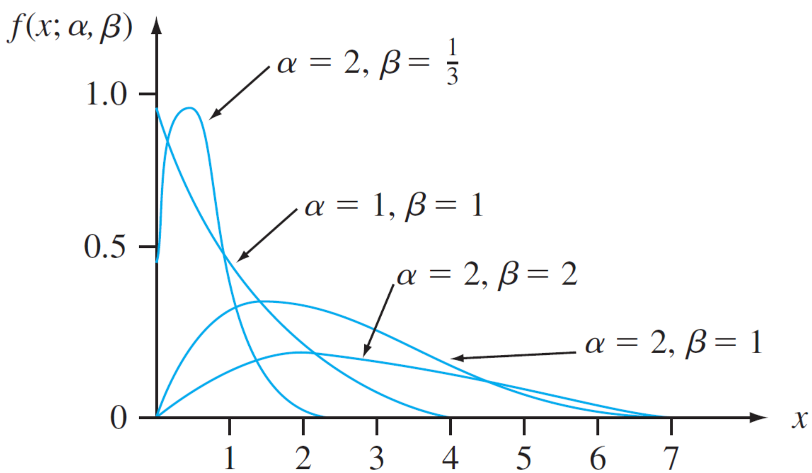
A continuous random variable X is said to have a **gamma distribution** if the pdf of X is

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-x/\beta} & x \geq 0 \\ 0 & \text{otherwise} \end{cases} \tag{4.8}$$

where the parameters α and β satisfy $\alpha > 0$, $\beta > 0$. The **standard gamma distribution** has $\beta = 1$, so the pdf of a standard gamma rv is given by (4.7).

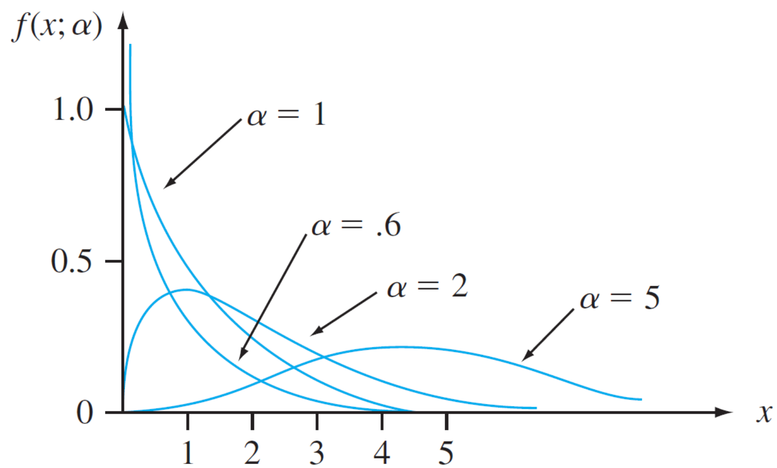
The Gamma Distribution

- The exponential distribution results from taking $\alpha = 1$ and $\beta = 1/\lambda$. Figure 4.27(a) illustrates the graphs of the gamma pdf $f(x; \alpha, \beta)$ (4.8) for several (α, β) pairs, whereas Figure 4.27(b) presents graphs of the standard gamma pdf.



Gamma density curves

Figure 4.27(a)



standard gamma density curves

Figure 4.27(b)

The Gamma Distribution

- For the standard pdf, when $\alpha \leq 1$, $f(x; \alpha)$, is strictly decreasing as x increases from 0; when $\alpha > 1$, $f(x; \alpha)$ rises from 0 at $x = 0$ to a maximum and then decreases.

- The parameter β in (4.8) is called the *scale parameter* because values other than 1 either stretch or compress the pdf in the x direction.

The Gamma Distribution

The mean and variance of a random variable *X* having the gamma distribution $f(x; \alpha, \beta)$ are

$$E(X) = \mu = \alpha\beta \qquad V(X) = \sigma^2 = \alpha\beta^2$$

When *X* is a standard gamma rv, the cdf of *X*,

$$F(x; \alpha) = \int_0^x \frac{y^{\alpha-1}e^{-y}}{\Gamma(\alpha)}\,dy \qquad x > 0$$

is called the **incomplete gamma function** [sometimes the incomplete gamma function refers to Expression (4.9) without the denominator $\Gamma(\alpha)$ in the integrand].

The Gamma Distribution

- There are extensive tables of available $F(x; \alpha)$; in Appendix Table A.4, we present a small tabulation for $\alpha = 1, 2, \ldots, 10$ and $x = 1, 2, \ldots, 15$.

Table A.4 The Incomplete Gamma Function

$$F(x; \alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y}\, dy$$

$x \backslash \alpha$	1	2	3	4	5	6	7	8	9	10
1	.632	.264	.080	.019	.004	.001	.000	.000	.000	.000
2	.865	.594	.323	.143	.053	.017	.005	.001	.000	.000
3	.950	.801	.577	.353	.185	.084	.034	.012	.004	.001
4	.982	.908	.762	.567	.371	.215	.111	.051	.021	.008
5	.993	.960	.875	.735	.560	.384	.238	.133	.068	.032
6	.998	.983	.938	.849	.715	.554	.394	.256	.153	.084
7	.999	.993	.970	.918	.827	.699	.550	.401	.271	.170
8	1.000	.997	.986	.958	.900	.809	.687	.547	.407	.283
9		.999	.994	.979	.945	.884	.793	.676	.544	.413
10		1.000	.997	.990	.971	.933	.870	.780	.667	.542
11			.999	.995	.985	.962	.921	.857	.768	.659
12			1.000	.998	.992	.980	.954	.911	.845	.758
13				.999	.996	.989	.974	.946	.900	.834
14				1.000	.998	.994	.986	.968	.938	.891
15					.999	.997	.992	.982	.963	.930

Appendix A23 (Walpole)

Table A.23 The Incomplete Gamma Function: $F(x;\alpha) = \int_0^x \frac{1}{\Gamma(\alpha)} y^{\alpha-1} e^{-y}\, dy$

	α									
x	1	2	3	4	5	6	7	8	9	10
1	0.6320	0.2640	0.0800	0.0190	0.0040	0.0010	0.0000	0.0000	0.0000	0.0000
2	0.8650	0.5940	0.3230	0.1430	0.0530	0.0170	0.0050	0.0010	0.0000	0.0000
3	0.9500	0.8010	0.5770	0.3530	0.1850	0.0840	0.0340	0.0120	0.0040	0.0010
4	0.9820	0.9080	0.7620	0.5670	0.3710	0.2150	0.1110	0.0510	0.0210	0.0080
5	0.9930	0.9600	0.8750	0.7350	0.5600	0.3840	0.2380	0.1330	0.0680	0.0320
6	0.9980	0.9830	0.9380	0.8490	0.7150	0.5540	0.3940	0.2560	0.1530	0.0840
7	0.9990	0.9930	0.9700	0.9180	0.8270	0.6990	0.5500	0.4010	0.2710	0.1700
8	1.0000	0.9970	0.9860	0.9580	0.9000	0.8090	0.6870	0.5470	0.4070	0.2830
9		0.9990	0.9940	0.9790	0.9450	0.8840	0.7930	0.6760	0.5440	0.4130
10		1.0000	0.9970	0.9900	0.9710	0.9330	0.8700	0.7800	0.6670	0.5420
11			0.9990	0.9950	0.9850	0.9620	0.9210	0.8570	0.7680	0.6590
12			1.0000	0.9980	0.9920	0.9800	0.9540	0.9110	0.8450	0.7580
13				0.9990	0.9960	0.9890	0.9740	0.9460	0.9000	0.8340
14				1.0000	0.9980	0.9940	0.9860	0.9680	0.9380	0.8910
15					0.9990	0.9970	0.9920	0.9820	0.9630	0.9300

Example 4.23

The article “The Probability Distribution of Maintenance Cost of a System Affected by the Gamma Process of Degradation” (*Reliability Engr. and System Safety*, 2012: 65–76) notes that the gamma distribution is widely used to model the extent of degradation such as corrosion, creep, or wear.

Let X represent the amount of degradation of a certain type, and suppose that it has a standard gamma distribution with $\alpha = 2$. Since

$$P(a \leq X \leq b) = F(b) - F(a)$$

Example 4.23

cont'd

When X is continuous,

$$P(3 \le X \le 5) = F(5;\, 2) - F(3;\, 2) = .960 - .801$$

$$= .159$$

The probability that the reaction time is more than 4 sec is

$$P(X > 4) = 1 - P(X \le 4) \quad = 1 - F(4;\, 2)$$

The Gamma Distribution

- The incomplete gamma function can also be used to compute probabilities involving nonstandard gamma distributions. These probabilities can also be obtained almost instantaneously from various software packages.
- **Proposition**

Let X have a gamma distribution with parameters α and β. Then for any $x > 0$, the cdf of X is given by

$$P(X \leq x) = F(x; \alpha, \beta) = F\left(\frac{x}{\beta}; \alpha\right)$$

where $F(\,\cdot\,; \alpha)$ is the incomplete gamma function.

Example 4.24

- Suppose the survival time X in weeks of a randomly selected male mouse exposed to 240 rads of gamma radiation has a gamma distribution with $\alpha = 8$ and $\beta = 15$.

- (Data in *Survival Distributions: Reliability Applications in the Biomedical Services,* by A. J. Gross and V. Clark, suggests $\alpha \approx 8.5$ and $\beta \approx 13.3$.)

- The expected survival time is $E(X) = (8)(5) = 120$ weeks, whereas $V(X) = (8)(15)^2 = 1800$ and $\sigma_X = \sqrt{1800} = 42.43$ weeks.

Example 4.24

cont'd

The probability that a mouse survives between 60 and 120 weeks is

$$P(60 \leq X \leq 120) = P(X \leq 120) - P(X \leq 60)$$

$$= F(120/15;\, 8) - F(60/15;\, 8)$$

$$= F(8;8) - F(4;8)$$

$$= .547 - .051$$

$$= .496$$

cont'd

Example 4.24

The probability that a mouse survives at least 30 weeks is

$$P(X \geq 30) = 1 - P(X < 30)$$

$$= 1 - P(X \leq 30)$$

$$= 1 - F(30/15;\ 8)$$

$$= .999$$

Erlang Distribution

- The gamma distribution for α an integer (also called the Erlang distribution) is shown below.

$$f(t) = \frac{\lambda(\lambda t)^{\alpha-1} e^{-\lambda t}}{(\alpha-1)!} \qquad for\ t>0$$

$$= \quad 0 \qquad\qquad elsewhere$$

Lognormal Distribution

- Lognormal distribution. For some applications, the natural log of the observations follows a normal distribution. Here μ and σ are specified in terms of the logs of the original data.
  - Then $z = (\ln(x) - \mu)/\sigma$ follows a standard normal distribution.
  - Working backwards, how do we convert back to x?
  - To convert back to x, we use $x = e^{\mu + z\sigma}$.
- The mean and the variance of the lognormal distribution (in the units of the original data) are:

$$E(X) = e^{\mu + \frac{\sigma^2}{2}} \qquad Var(X) = e^{2\mu + \sigma^2}(e^{\sigma^2} - 1)$$

The Lognormal Distribution

- **Definition**

A nonnegative rv X is said to have a **lognormal distribution** if the rv $Y = \ln(X)$ has a normal distribution. The resulting pdf of a lognormal rv when $\ln(X)$ is normally distributed with parameters μ and σ is

$$f(x;\mu,\sigma) = \begin{cases} \dfrac{1}{\sqrt{2\pi}\sigma x} e^{-[\ln(x)-\mu]^2/(2\sigma^2)} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

The Lognormal Distribution

- Be careful here; the parameters μ and σ are not the mean and standard deviation of *X* but of ln(*X*).

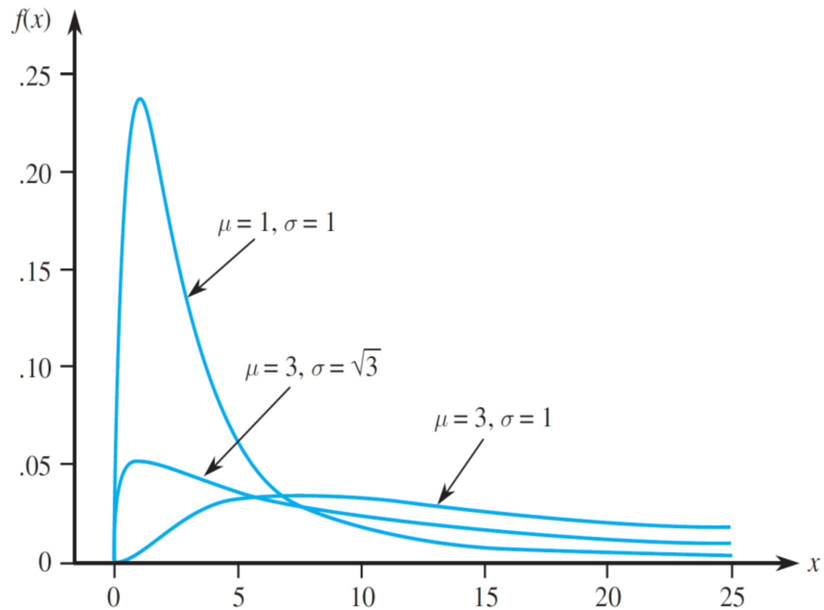
  It is common to refer to μ and σ as the location and the scale parameters, respectively. The mean and variance of *X* can be shown to be

$$E(X) = e^{\mu+\sigma^2/2} \qquad V(X) = e^{2\mu+\sigma^2} \cdot (e^{\sigma^2} - 1)$$

- In Chapter 5, we will present a theoretical justification for this distribution in connection with the Central Limit Theorem, but as with other distributions, the lognormal can be used as a model even in the absence of such justification.

The Lognormal Distribution

- Figure 4.30 illustrates graphs of the lognormal pdf; although a normal curve is symmetric, a lognormal curve has a positive skew.



Lognormal density curves

Figure 4.30

The Lognormal Distribution

- Because ln(X) has a normal distribution, the cdf of X can be expressed in terms of the cdf $\phi(z)$ of a standard normal rv Z.

$$
\begin{aligned}
F(x;\mu,\sigma) &= P(X \le x) = P[\ln(X) \le \ln(x)] \\
&= P\left(Z \le \frac{\ln(x)-\mu}{\sigma}\right) = \Phi\left(\frac{\ln(x)-\mu}{\sigma}\right) \qquad x \ge 0 \qquad (4.13)
\end{aligned}
$$

Example 4.27

- According to the article "Predictive Model for Pitting Corrosion in Buried Oil and Gas Pipelines" (*Corrosion*, 2009: 332–342), the lognormal distribution has been reported as the best option for describing the distribution of maximum pit depth data from cast iron pipes in soil.

- The authors suggest that a lognormal distribution with $\mu = .353$ and $\sigma = .754$ is appropriate for maximum pit depth (mm) of buried pipelines.

- For this distribution, the mean value and variance of pit depth are

$$E(X) = e^{.353+(.754)^2/2} = e^{.6373} = 1.891$$

Example 4.27

cont'd

$$V(X) = e^{2(.353)+(.754)^2} \cdot (e^{(.754)^2} - 1) = (3.57697)(.765645) = 2.7387$$

The probability that maximum pit depth is between 1 and 2 mm is

$$P(1 \le X \le 2) = P(\ln(1) \le \ln(X) \le \ln(2))$$

$$= P(0 \le \ln(X) \le .693)$$

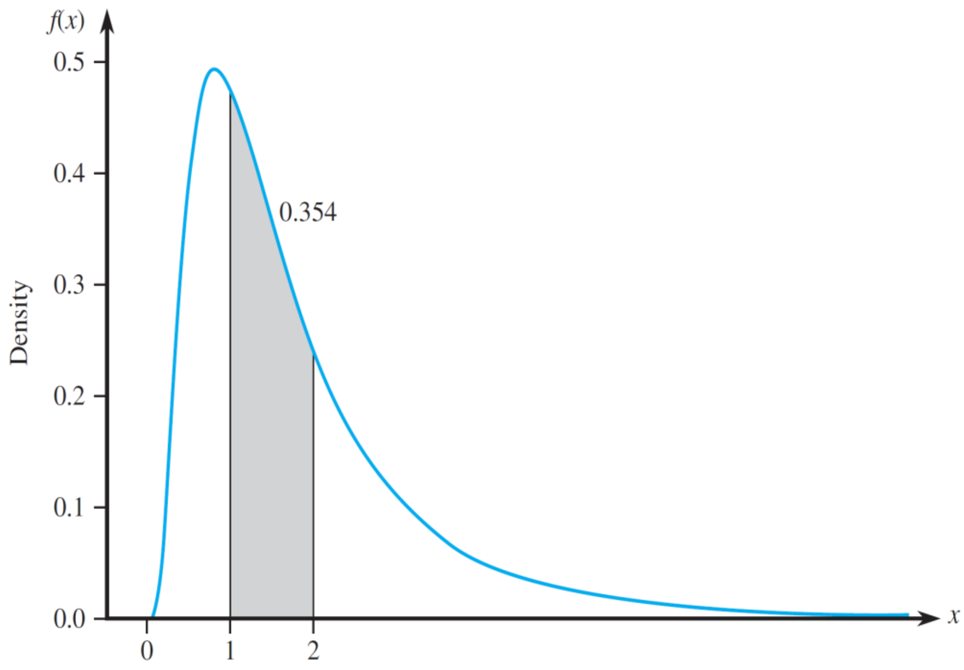
$$= P\left(\frac{0 - .353}{.754} \le Z \le \frac{.693 - .353}{.754}\right)$$

$$= \phi(.47) - \phi(-.45) = .354$$

Example 4.27

cont'd

- This probability is illustrated in Figure 4.31 (from Minitab).



Lognormal density curve with location = .353 and scale = .754

Figure 4.31

Example 4.27

cont'd

- What value c is such that only 1% of all specimens have a maximum pit depth exceeding c? The desired value satisfies

$$.99 = P(X \le c) = P\left(Z \le \frac{\ln(c) - .353}{.754}\right)$$

- The z critical value 2.33 captures an upper-tail area of .01 ($z_{.01}$ = 2.33), and thus a cumulative area of .99.

This implies th $\dfrac{\ln(c) - .353}{.754} = 2.33$

Example 4.27

cont'd

- From which $\ln(c) = 2.1098$ and $c = 8.247$.

- Thus 8.247 is the 99th percentile of the maximum pit depth distribution.

Other Continuous Distributions

- Chi-Squared distribution. Used for statistical inference. We will use the chi-squared table in the appendix later in the semester.
- Weibull distribution. Has many applications in reliability and life testing and is used to model failure rate.
  - The failure rate at time t for the Weibull distribution is

$$Z(t) = \alpha\beta t^{\beta-1}, \qquad for\ t > 0$$

  - For $\beta = 1$ failure rate is constant (memoryless).
  - For $\beta > 1$ failure rate increases over time.
  - For $\beta < 1$ failure rate decreases over time.

The Chi-Squared Distribution

- The chi-squared distribution is important because it is the basis for a number of procedures in statistical inference.

- The central role played by the chi-squared distribution in inference springs from its relationship to normal distributions.

The Chi-Squared Distribution

- **Definition**

Let ν be a positive integer. Then a random variable X is said to have a **chi-squared distribution** with parameter ν if the pdf of X is the gamma density with $\alpha = \nu/2$ and $\beta = 2$. The pdf of a chi-squared rv is thus

$$f(x;\nu) = \begin{cases} \dfrac{1}{2^{\nu/2}\Gamma(\nu/2)} x^{(\nu/2)-1} e^{-x/2} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad (4.10)$$

The parameter ν is called the **number of degrees of freedom** (df) of X. The symbol χ^2 is often used in place of "chi-squared."

The Weibull Distribution

- The family of Weibull distributions was introduced by the Swedish physicist Waloddi Weibull in 1939; his 1951 article "A Statistical Distribution Function of Wide Applicability" (*J. of Applied Mechanics*, vol. 18: 293–297) discusses a number of applications.

Definition

A random variable X is said to have a **Weibull distribution** with shape parameter α and scale parameter β ($\alpha > 0$, $\beta > 0$) if the pdf of X is

$$f(x; \alpha, \beta) = \begin{cases} \dfrac{\alpha}{\beta^\alpha} x^{\alpha-1} e^{-(x/\beta)^\alpha} & x \geq 0 \\ 0 & x < 0 \end{cases} \tag{4.11}$$

The Weibull Distribution

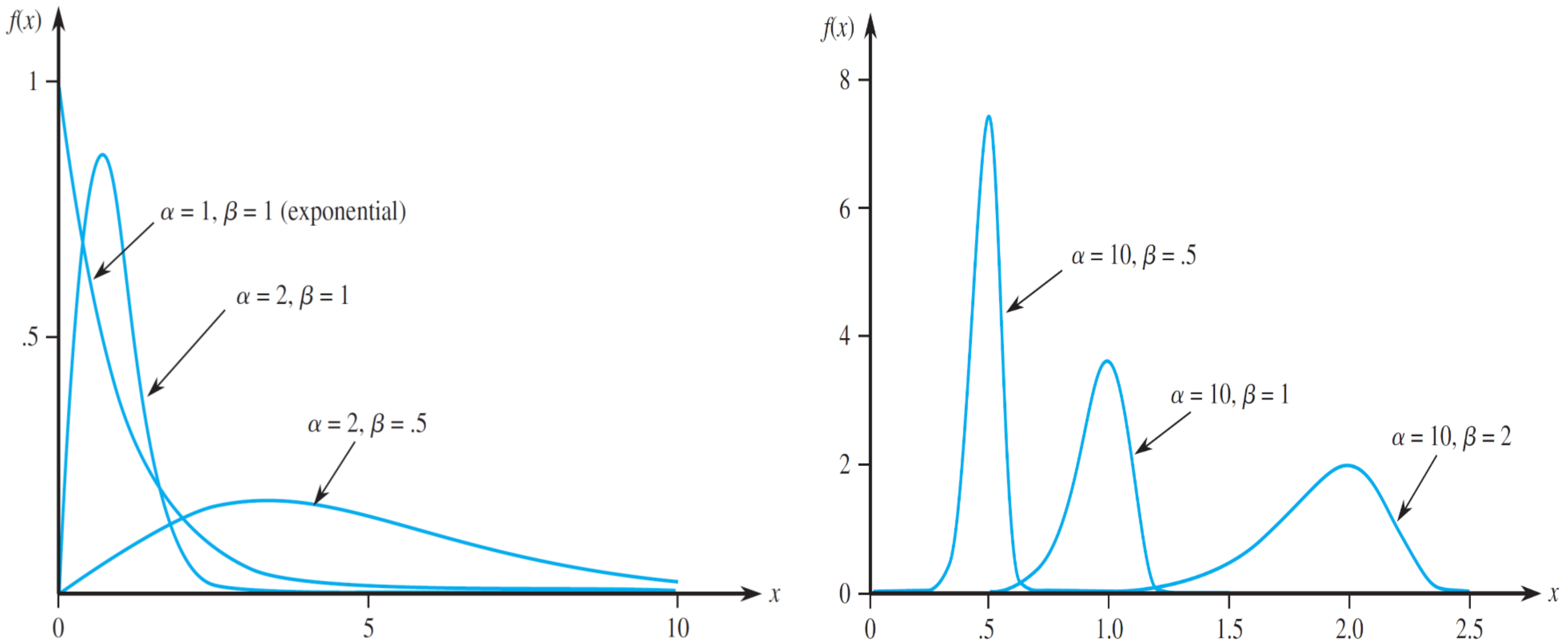
- In some situations, there are theoretical justifications for the appropriateness of the Weibull distribution, but in many applications $f(x; \alpha, \beta)$ simply provides a good fit to observed data for particular values of α and β.

  When $\alpha = 1$, the pdf reduces to the exponential distribution (with $\lambda = 1/\beta$), so the exponential distribution is a special case of both the gamma and Weibull distributions.

  However, there are gamma distributions that are not Weibull distributions and vice versa, so one family is not a subset of the other.

The Weibull Distribution

- Both α and β can be varied to obtain a number of different-looking density curves, as illustrated in Figure 4.28.



Weibull density curves

Figure 4.28

The Weibull Distribution

- β is called a scale parameter, since different values stretch or compress the graph in the x direction, and α is referred to as a shape parameter.

Integrating to obtain $E(X)$ and $E(X^2)$ yields

$$\mu = \beta\Gamma\left(1 + \frac{1}{\alpha}\right) \qquad \sigma^2 = \beta^2\left\{\Gamma\left(1 + \frac{2}{\alpha}\right) - \left[\Gamma\left(1 + \frac{1}{\alpha}\right)\right]^2\right\}$$

The computation of μ and σ^2 thus necessitates using the gamma function.

The Weibull Distribution

- The integration $\int_0^x f(y;\alpha,\beta)\,dy$ is easily carried out to obtain the cdf of *X*.

The cdf of a Weibull rv having parameters α and β is

$$F(x;\alpha,\beta) = \begin{cases} 0 & x < 0 \\ 1 - e^{-(x/\beta)^\alpha} & x \ge 0 \end{cases} \tag{4.12}$$

Example 4.25

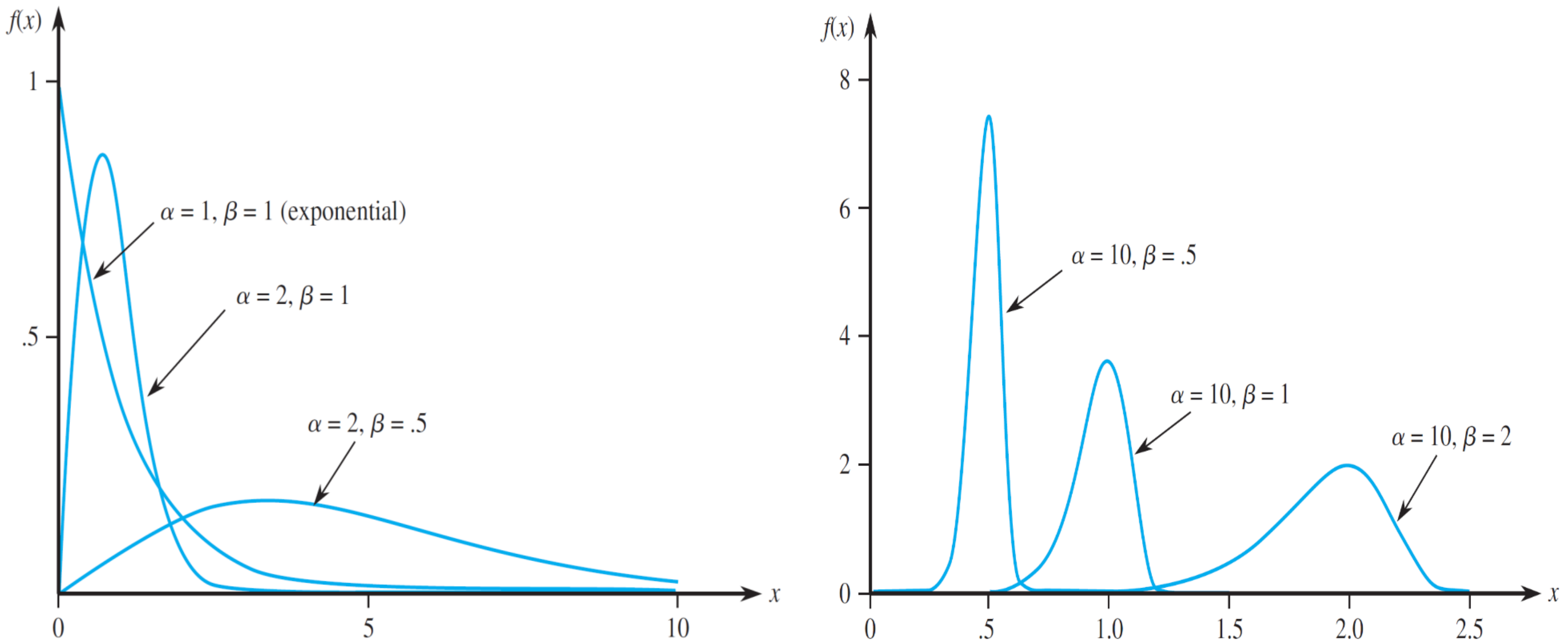
- In recent years the Weibull distribution has been used to model engine emissions of various pollutants.

- Let X denote the amount of NO_x emission (g/gal) from a randomly selected four-stroke engine of a certain type, and suppose that X has a Weibull distribution with $\alpha = 2$ and $\beta = 10$ (suggested by information in the article "Quantification of Variability and Uncertainty in Lawn and Garden Equipment NO_x and Total Hydrocarbon Emission Factors," *J. of the Air and Waste Management Assoc.*, 2002: 435–448).

cont’d

Example 4.25

- The corresponding density curve looks exactly like the one in Figure 4.28 for $\alpha = 2$, $\beta = 1$



Weibull density curves

Figure 4.28

Example 4.25

cont'd

Except that now the values 50 and 100 replace 5 and 10 on the horizontal axis (because β is a "scale parameter").

Then

$$
\begin{aligned}
P(X \leq 10) &= F(10; 2, 10) \\
&= 1 - e^{-(10/10)^2} \\
&= 1 - e^{-1} \\
&= .632
\end{aligned}
$$

Similarly, $P(X \leq 25) = .998$, so the distribution is almost entirely concentrated on values between 0 and 25.

Example 4.25

cont'd

- The value c which separates the 5% of all engines having the largest amounts of NO_x emissions from the remaining 95% satisfies

$$.95 = 1 - e^{-(c/10)^2}$$

- Isolating the exponential term on one side, taking logarithms, and solving the resulting equation gives $c \approx 17.3$ as the 95th percentile of the emission distribution.

Thank You

"We trust in GOD, all others must bring data"