The Normal Probability Distribution

Chapter 7

McGraw-Hill/Irwin

連續隨機變數 vs. 離散隨機變數 (1)

- 假設在 {2, 3, 4, 5} 內隨機抽取一數字，每個數字被抽到的機率 0.25。
- 假設在區間 [2, 5] 內，隨機抽取一個數字，則每個數字被抽到的機率為何？
  – 由於區間 [2, 5] 內的數字有無限多個，所以任意一點被抽到的機率為 0 (= 1/∞)。

2

連續隨機變數 vs. 離散隨機變數 (2)

- 離散隨機變數與連續隨機變數最大的差別在於：
  – 離散隨機變數，單點發生的機率是大於等於 0。
  – 連續隨機變數單一點發生機率等於0。
- 描述離散隨機變數的機率分配，只要描述每個可能點的機率即可。

$$P(X = x) = 1/4, \quad x = 2,3,4,5$$

3

連續隨機變數 vs. 離散隨機變數 (3)

- 由於連續隨機變數可能發生點的機率均為 0，因此無法描述機率分配，改以描述每個點可能發生的機率密度 (Probability Density)。

$$f(x) = 1/3, \quad 2 \le x \le 5$$

  – f(x = 2) =1/ 3，並不表示隨機變數 X = 2 的機率是 1/3，僅表示 X=2 這點的機率密度等於 1/3。
  – **質量=密度×體積 vs. 機率值 = 機率密度×區間**
    在連續機率分配，不求單點發生機率，只求區間發生機率，因此 X 發生在區間 [2, 3.5]內的機率等於 0.5。

$$P(2 \le x \le 3.5) = \sum_{x=2}^{3.5} f(x) = \sum_{x=2}^{3.5} \frac{1}{3} = \frac{3.5-2}{5-2} = 0.5$$

4

連續隨機變數 vs. 離散隨機變數 (4)

- 連續隨機變數的機率密度 (Probability Density) 函數，值可能大於 1。

$$f(x)=3, \quad 0 \le x \le \frac{1}{3}$$

– f(x = 0.25) = 3，並不表示隨機變數 X = 0.25 的機率是 3，僅表示 X= 0.25 這點的機率密度等於 3。

5

Example：連續隨機變數

- 假設某股票的平均報酬率 5%，報酬率標準差 30 %。假設今日股價 $10，則一年後的股價 S 的機率分配為：

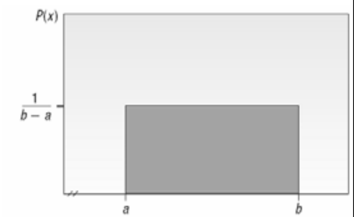
$$P(S=s)=\frac{1}{0.3S\sqrt{2\pi}}e^{-\frac{1}{2}\left(\frac{\ln S-\ln 10-0.05}{0.3}\right)^2}, \quad 0<S<\infty$$

6

Uniform Distribution (均勻分配)

- 假設在區間 [a, b] 內，隨機抽取一個數字 X，則每個數字被抽到的機率密度為何？一般用均勻分配來描述抽到的機率密度。
- 隨機變數 X 服從 Uniform Distribution 以 U [a,b] 表示，X 的機率密度函數為：

$$f(x)=\frac{1}{b-a}, \quad a \le x \le b$$



7

The Uniform Distribution – Mean and Standard Deviation

MEAN OF THE UNIFORM DISTRIBUTION $\mu=\frac{a+b}{2}$ [7–1]

STANDARD DEVIATION OF THE UNIFORM DISTRIBUTION $\sigma=\sqrt{\frac{(b-a)^2}{12}}$ [7–2]

UNIFORM DISTRIBUTION $P(x)=\frac{1}{b-a}$ if $a \le x \le b$ and 0 elsewhere [7–3]

8

The Uniform Distribution - Example

Southwest Arizona State University provides bus service to students while they are on campus. A bus arrives at the North Main Street and College Drive stop every 30 minutes between 6 A.M. and 11 P.M. during weekdays. Students arrive at the bus stop at random times. The time that a student waits is uniformly distributed from 0 to 30 minutes.

1. Draw a graph of this distribution.

2. How long will a student "typically" have to wait for a bus? In other words what is the mean waiting time? What is the standard deviation of the waiting times?

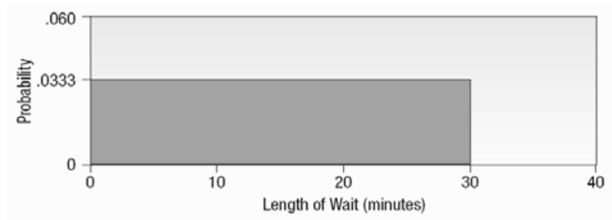
3. What is the probability a student will wait more than 25 minutes?

4. What is the probability a student will wait between 10 and 20 minutes?

9

The Uniform Distribution - Example

Draw a graph of this distribution.



10

The Uniform Distribution - Example

How long will a student "typically" have to wait for a bus? In other words what is the mean waiting time? What is the standard deviation of the waiting times?

$$\mu = \frac{a+b}{2} = \frac{0+30}{2} = 15$$

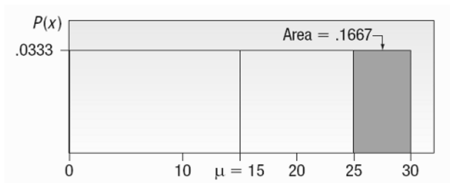
$$\sigma = \sqrt{\frac{(b-a)^2}{12}} = \sqrt{\frac{(30-0)^2}{12}} = 8.66$$

11

The Uniform Distribution - Example

What is the probability a student will wait more than 25 minutes?

$$P(25 < \text{wait time} < 30) = (\text{height})(\text{base})$$
$$= \frac{1}{(30-0)}(5) = 0.1667$$

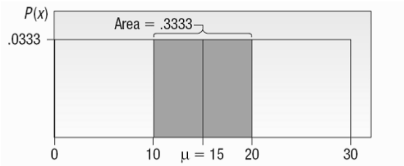


12

The Uniform Distribution - Example

What is the probability a student will wait between 10 and 20 minutes?

$$P(10 < \text{wait time} < 20) = (\text{height})(\text{base})$$
$$= \frac{1}{(30-0)}(10) = 0.3333$$



13

常態分配 (Normal Distribution)

- 一般來說，一個常態班級的考試成績，大多以平均分數為中心，呈鐘型分配。身高，體重、年收入、、、等，大多有類似的性質，我們用常態分配來描述這類型的資料。

14

常態分配 (Normal Distribution)

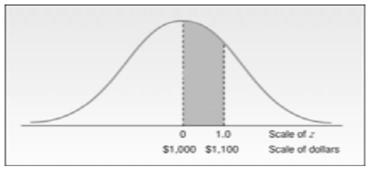
假設隨機變數 X 服從常態分配，平均數與變異數分別爲 $\mu = E[X]$ 與 $\sigma^2 = E[(X-\mu)^2]$，一般以 $X \sim N(\mu, \sigma^2)$ 表示。則隨機變數 X 的機率密度函數：

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

15

Characteristics of a Normal Probability Distribution

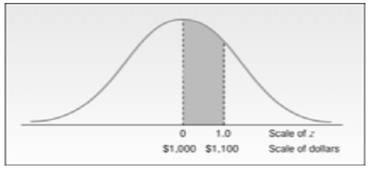
- It is **bell-shaped** and has a single peak at the center of the distribution.
- The arithmetic mean, median, and mode are equal.
- The total area under the curve is 1; half the area under the normal curve is to the right of this center point and the other half to the left of it.
- It is **symmetrical** about the mean.



16

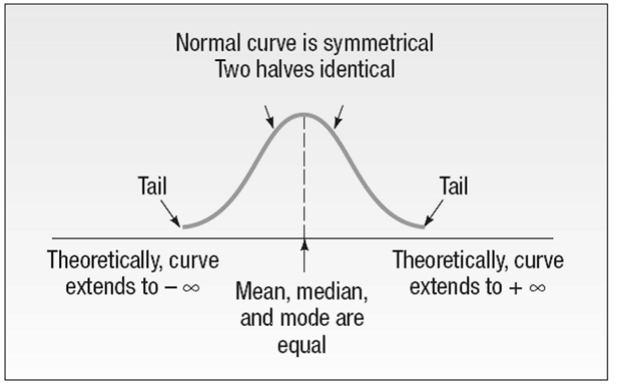
Characteristics of a Normal Probability Distribution

- It is **asymptotic:** The curve gets closer and closer to the *X*-axis but never actually touches it. To put it another way, the tails of the curve extend indefinitely in both directions.
- The location of a normal distribution is determined by the mean, μ, the dispersion or spread of the distribution is determined by the standard deviation, σ .



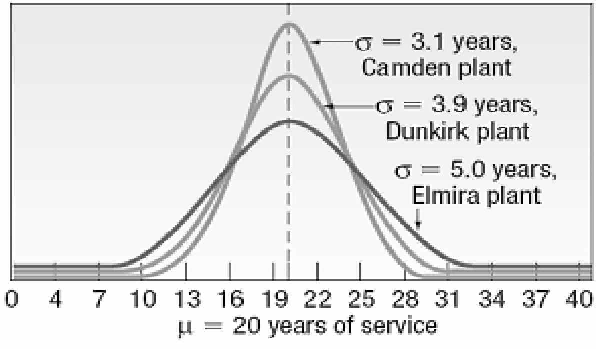
17

The Normal Distribution - Graphically



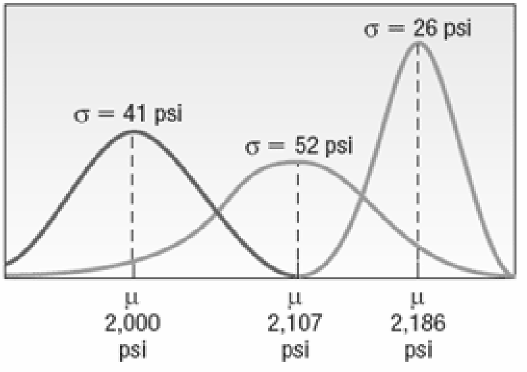
18

The Normal Distribution - Families



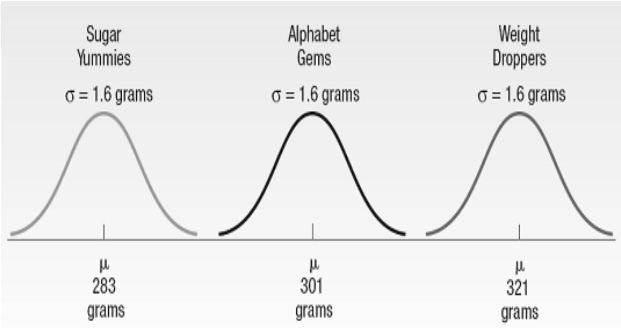
19

The Normal Distribution - Families



20

The Normal Distribution - Families



21

The Standard Normal Probability Distribution (標準常態分配)

- The standard normal distribution is a normal distribution with $\mu = 0$ and $\sigma^2 = 1$.
- It is also called the *Z* distribution.
- A *Z*-value is the distance between a selected value *X* and the population mean μ, divided by the population standard deviation σ.
- The formula is:

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

22

常態分配的線性轉換

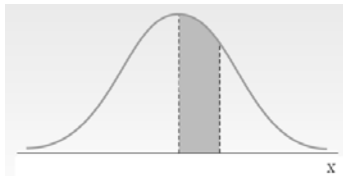
- 若 a 與 b 是常數且 $X \sim N(\mu, \sigma^2)$，令 $Y = a + bX$，則 $Y \sim N(a + b\mu, b^2\sigma^2)$

- 若 $X \sim N(\mu, \sigma^2)$，令 $Z = (X - \mu) / \sigma$，則 $Z \sim N(0, 1)$，標準常態分配。

23

The Normal Probability Distribution (常態分配)

- 任何常態分配 $X \sim N(\mu, \sigma^2)$：
  - $P(X < 0)$ 表示 X 值小於 0 的機率。
  - $P(X < 2.7)$ 表示 X 值小於 2.7 的機率。
  - $P(X < -3)$ 表示 X 值小於 - 3 的機率。



24

The Normal Probability Distribution (常態分配)

- 任意一常態隨機變數 X ～ N(μ, σ^2) 並給定一任意常數 a，則 P(X < a) 的機率，僅與 a 和 μ 相距幾個 σ 單位有關。
- 例如：已知 X ～ N(10, 100) 且 Y ～ N(-8, 64)，則:
  - P(X < 20) = P(Y< 0)
  - P(X<30) =P(Y < 8)
  - P(X > 0) = P(Y > -16)
  - P(X > -10) = P(Y > -24)

The Normal Probability Distribution (常態分配)

- 任意一常態隨機變數 X ～ N(μ, σ^2) 並給定一任意常數 a，則 P(X < a) 的機率，僅與 a 和 μ 相距幾個 σ 單位有關。因此：

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

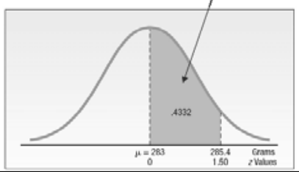
- 例如：已知 X ～ N(10, 100) 且 Y ～ N(-8, 64)，則:
  - P(X < 20) = P(Y< 0) = P(Z < 1)
  - P(X<30) =P(Y < 8) = P(Z < 2)
  - P(X > 0) = P(Y > -16) = P(Z > -1)
  - P(X > -10) = P(Y > -24) = P(Z > -2)

The Normal Probability Distribution (常態分配)

- 任意一常態隨機變數 X ～ N(μ, σ^2) 並給定一任意常數 a，則求算 P(X < a) 的機率即等於求算 P(Z < (a-μ) / σ) 的機率。
- 要求算常態隨機變數 X ～ N(μ, σ^2) 的相關機率，皆可以轉換成標準常態分配 Z ～ N(0, 1) 的相關機率。
- 透過標準常態分配表，常態隨機變數 X ～ N(μ, σ^2) 的相關機率，皆可以查表求得。
- 標準常態分配表見課本 Appendix B.1, PP. 784.

Areas Under the Normal Curve

z	0.00	0.01	0.02	0.03	0.04	0.05	. . .
1.3	0.4032	0.4049	0.4066	0.4082	0.4099	0.4115	
1.4	0.4192	0.4207	0.4222	0.4236	0.4251	0.4265	
1.5	0.4332	0.4345	0.4357	0.4370	0.4382	0.4394	
1.6	0.4452	0.4463	0.4474	0.4484	0.4495	0.4505	
1.7	0.4554	0.4564	0.4573	0.4582	0.4591	0.4599	
1.8	0.4641	0.4649	0.4656	0.4664	0.4671	0.4678	
1.9	0.4713	0.4719	0.4726	0.4732	0.4738	0.4744	
.							
.							
.							



The Normal Distribution – Example

The weekly incomes of shift foremen in the glass industry follow the normal probability distribution with a mean of $1,000 and a standard deviation of $100. What is the *z* value for the income, let's call it *X,* of a foreman who earns $1,100 per week? For a foreman who earns $900 per week?

For X = $1,100:

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}100 - \$1{,}000}{\$100} = 1.00$$

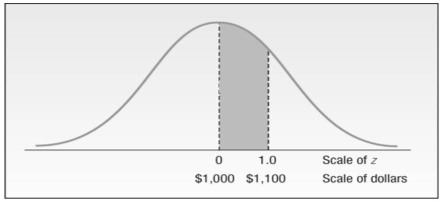
For X = $900:

$$z = \frac{X - \mu}{\sigma} = \frac{\$900 - \$1{,}000}{\$100} = -1.00$$

Normal Distribution – Finding Probabilities

In an earlier example we reported that the mean weekly income of a shift foreman in the glass industry is normally distributed with a mean of $1,000 and a standard deviation of $100.

What is the likelihood of selecting a foreman whose weekly income is between $1,000 and $1,100?



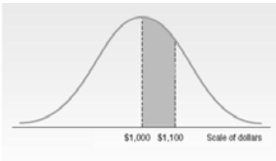
Normal Distribution – Finding Probabilities

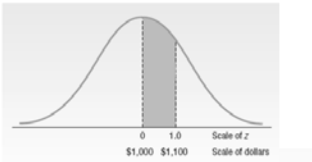
For X = $1,000:

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}000 - \$1{,}000}{\$100} = 0.00$$

For X = $1,100:

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}100 - \$1{,}000}{\$100} = 1.00$$





Normal Distribution – Finding Probabilities (Example 2)

Refer to the information regarding the weekly income of shift foremen in the glass industry. The distribution of weekly incomes follows the normal probability distribution with a mean of $1,000 and a standard deviation of $100.

What is the probability of selecting a shift foreman in the glass industry whose income is:

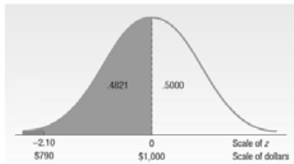
Between $790 and $1,000?

For X = $790:

$$z = \frac{X - \mu}{\sigma} = \frac{\$790 - \$1{,}000}{\$100} = -2.10$$

For X = $1,000:

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}000 - \$1{,}000}{\$100} = 0.00$$



Normal Distribution – Finding Probabilities (Example 3)

Refer to the information regarding the weekly income of shift foremen in the glass industry. The distribution of weekly incomes follows the normal probability distribution with a mean of $1,000 and a standard deviation of $100.

What is the probability of selecting a shift foreman in the glass industry whose income is:

Less than $790?

Find Z for X = $790 :

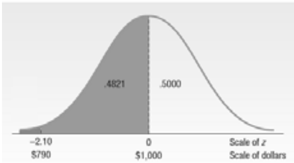
$$z = \frac{X - \mu}{\sigma} = \frac{\$790 - \$1{,}000}{\$100} = -2.10$$

To find the are below - 2.10,

subtract from 0.50 the area from - 2.10 to 0

= 0.50-0.4821

= 0.0179



33

Normal Distribution – Finding Probabilities (Example 4)

Refer to the information regarding the weekly income of shift foremen in the glass industry. The distribution of weekly incomes follows the normal probability distribution with a mean of $1,000 and a standard deviation of $100.

What is the probability of selecting a shift foreman in the glass industry whose income is:

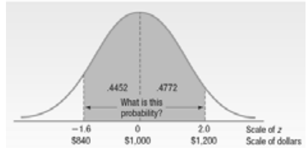
Between $840 and $1,200?

For X = $840:

$$z = \frac{X - \mu}{\sigma} = \frac{\$840 - \$1{,}000}{\$100} = -1.60$$

For X = $1,200:

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}200 - \$1{,}000}{\$100} = 2.00$$



34

Normal Distribution – Finding Probabilities (Example 5)

Refer to the information regarding the weekly income of shift foremen in the glass industry. The distribution of weekly incomes follows the normal probability distribution with a mean of $1,000 and a standard deviation of $100.

What is the probability of selecting a shift foreman in the glass industry whose income is:

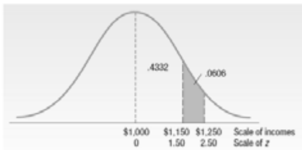
Between $1,150 and $1,250

For X = $1,150 :

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}150 - \$1{,}000}{\$100} = 1.50$$

For X = $1,250 :

$$z = \frac{X - \mu}{\sigma} = \frac{\$1{,}250 - \$1{,}000}{\$100} = 2.50$$



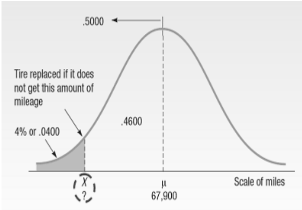
35

Using Z in Finding X Given Area - Example

Layton Tire and Rubber Company wishes to set a minimum mileage guarantee on its new MX100 tire. Tests reveal the mean mileage is 67,900 with a standard deviation of 2,050 miles and that the distribution of miles follows the normal probability distribution. It wants to set the minimum guaranteed mileage so that no more than 4 percent of the tires will have to be replaced. What minimum guaranteed mileage should Layton announce?

36

Using Z in Finding X Given Area - Example



Solve X using the formula :

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

$z = \frac{X - 67{,}900}{2{,}050}$

The value of z is found using the 4% information

The area between 67,900 and X is .4600, found by .5000 - .0400.

Using Appendix D, the area closest to .4600 is .4599, which gives a z value of 1.75.

$1.75 = \frac{X - 67{,}900}{2{,}050}$ then solving for X

$1.75(2{,}050) = X - 67{,}900$

$X = 67{,}900 - 1.75(2{,}050)$

$X = 64{,}312$

37

標準常態分配性質

- 標準常態分配 $Z \sim N(0,1)$ 以 0 為中心對稱。
  - $P(Z > a) = P(Z < -a)$
  - $P(Z > -a) = P(Z < a)$
  - $P(Z > a) = 1- P(Z < a) = 1- P(Z > -a)$
- 例如：
  - $P(Z < -0.5)$, $P(Z >2)$, $P(-1<Z<1.5)$
- z_α 的定義為：$P(Z > Z_\alpha) = \alpha$ 。
- $Z_\alpha = - Z_{1-\alpha}$
  - $Z_{0.025} = 1.96$
  - $Z_{0.05} = 1.645$
  - $Z_{0.95} = -1.645$

38

獨立常態分配之加法性

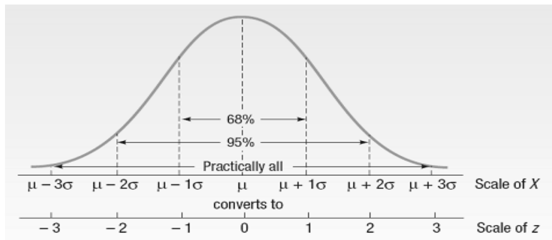
若 X_i, i=1,2,...., n，是獨立的常態分配，平均數 μ_i 且變異數 σ_i^2 。令 $a_1, a_2, ..., a_n$ 是常數且 $Y = a_1X_1 + a_2X_2 + ... + a_nX_n$，

則 $Y \sim N\left(\sum_{i=1}^{n} a_i\mu_i, \sum_{i=1}^{n} a_i^2\sigma_i^2\right)$

39

The Empirical Rule

- About 68 percent of the area under the normal curve is within one standard deviation of the mean.
- About 95 percent is within two standard deviations of the mean.
- Practically all is within three standard deviations of the mean.



40

The Empirical Rule - Example

As part of its quality assurance program, the Autolite Battery Company conducts tests on battery life. For a particular D-cell alkaline battery, the mean life is 19 hours. The useful life of the battery follows a normal distribution with a standard deviation of 1.2 hours.

Answer the following questions.

1. About 68 percent of the batteries failed between what two values?
2. About 95 percent of the batteries failed between what two values?
3. Virtually all of the batteries failed between what two values?

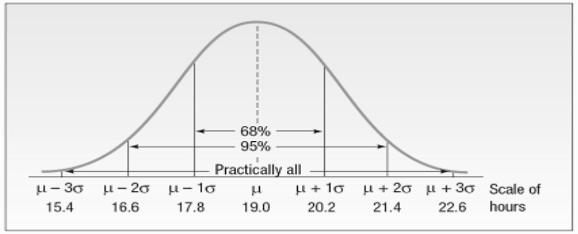
41

The Empirical Rule - Example

We can use the results of the Empirical Rule to answer these questions.

1. About 68 percent of the batteries will fail between 17.8 and 20.2 hours by 19.0 ± 1(1.2) hours.
2. About 95 percent of the batteries will fail between 16.6 and 21.4 hours by 19.0 ± 2(1.2) hours.
3. Virtually all failed between 15.4 and 22.6 hours, found by 19.0 ± 3(1.2)

This information is summarized on the following chart.



42

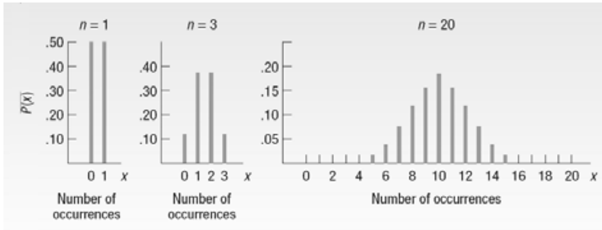
Normal Approximation to the Binomial

- The normal distribution (a continuous distribution) yields a good approximation of the binomial distribution (a discrete distribution) for large values of *n.*
- The normal probability distribution is generally a good approximation to the binomial probability distribution when $n\pi$ and $n(1-\pi)$ are both greater than 5.

43

Normal Approximation to the Binomial

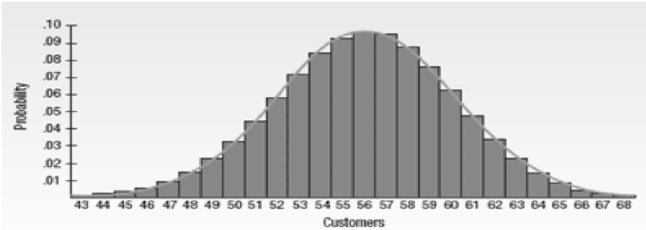
Using the normal distribution (a continuous distribution) as a substitute for a binomial distribution (a discrete distribution) for large values of *n* seems reasonable because, as *n* increases, a binomial distribution gets closer and closer to a normal distribution.



44

Continuity Correction Factor

The value .5 subtracted or added, depending on the problem, to a selected value when a binomial probability distribution (a discrete probability distribution) is being approximated by a continuous probability distribution (the normal distribution).



How to Apply the Correction Factor

Only four cases may arise. These cases are:

1. For the probability at least X occurs, use the area above (X -0.5).
2. For the probability that more than X occurs, use the area above (X+0.5).
3. For the probability that X or fewer occurs, use the area below (X -0.5).
4. For the probability that fewer than X occurs, use the area below (X+0.5).

Normal Approximation to the Binomial - Example

Suppose the management of the Santoni Pizza Restaurant found that 70 percent of its new customers return for another meal. For a week in which 80 new (first-time) customers dined at Santoni's, what is the probability that 60 or more will return for another meal?

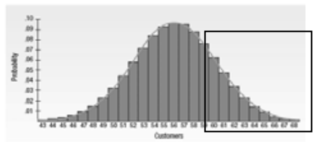
Normal Approximation to the Binomial - Example

$$P(x) = {}_{n}C_{x} (\pi)^{x} (1 - \pi)^{n-x}$$

$$P(x = 60) = {}_{80}C_{60} (.70)^{60} (1 - .70)^{20} = .063$$

$$P(x = 61) = {}_{80}C_{61} (.70)^{61} (1 - .70)^{19} = .048$$

Number Returning	Probability	Number Returning	Probability
43	.001	56	.097
44	.002	57	.095
45	.003	58	.088
46	.006	59	.077
47	.009	60	.063
48	.015	61	.048
49	.023	62	.034
50	.033	63	.023
51	.045	64	.014
52	.059	65	.008
53	.072	66	.004
54	.084	67	.002
55	.093	68	.001



$P(X \geq 60) = 0.063+0.048+ \ldots + 0.001) = 0.197$

Normal Approximation to the Binomial - Example

Step 1. Find the mean and the variance of a binomial distribution and find the *z* corresponding to an *X* of 59.5 (x-.5, the correction factor)

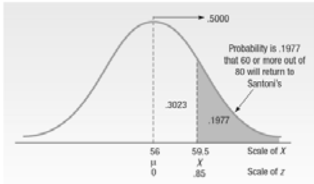
Step 2: Determine the area from 59.5 and beyond

$$\mu = n\pi = 80(.70) = 56$$

$$\sigma^2 = n\pi(1-\pi) = 80(.70)(1-.70) = 16.8$$

$$\sigma = \sqrt{16.8} = 4.10$$

$$z = \frac{X-\mu}{\sigma} = \frac{59.5-56}{4.10} = 0.85$$



Exercises

- 1,3,7,9,13,15,17,19,21,23,25,27,31,33,37,39, 41,43,45,53,55,59,63

End of Chapter 7