Discrete Probability Distributions

Discrete Uniform Probability Distribution:

- When the possible values of the probability mass function are all equal, then the probability distribution is a discrete uniform probability distribution.

DISCRETE UNIFORM PROBABILITY MASS FUNCTION

$$f(x) = 1/n \qquad \textbf{(4.15)}$$

- Where n = the number of unique values that may be assumed by the random variable.
- Example: Fair 6-sided die
  - Lottery – Probability for picking any one number is the same for all outcomes.

Discrete Probability Distributions

- Bernoulli Trial
  - An independent trial in which there are 2 outcomes
- Example:
  - Martin's an online specialty clothing store
  - Sends out targeted e-mails to its best customers
  - Special discounts available only to the recipients.
  - Each email is a Bernoulli Trial
    - Click Link = Success
    - Do not Click = Faiure

Discrete Probability Distributions

Binomial Probability Distribution:

- A **binomial probability distribution**
  - used to describe many situations in which a fixed number (*n*) of repeated identical and independent trials has two, and only two, possible outcomes:
  - Used to calculate the probability of a given number of successes in a set of (n) Bernoulli trials

Discrete Probability Distributions

The probability mass function for a binomial random variable that calculates the probability of *x* successes in *n* independent events.

BINOMIAL PROBABILITY MASS FUNCTION

$$f(x) = \binom{n}{x} p^x (1-p)^{(n-x)} \tag{4.16}$$

where

x = the number of successes
p = the probability of a success on one trial
n = the number of trials
$f(x)$ = the probability of x successes in n trials

and

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

Discrete Probability Distributions

Probability Distribution for the Number of Customers Who Click on the Link in the Martin's Targeted E-Mail

If you Send n = 3 Emails

x	f(x)
0	$\frac{3!}{0!3!}(0.30)^0(0.70)^3 = 0.343$
1	$\frac{3!}{1!2!}(0.30)^1(0.70)^2 = 0.441$
2	$\frac{3!}{2!1!}(0.30)^2(0.70)^1 = 0.189$
3	$\frac{3!}{3!0!}(0.30)^3(0.70)^0 = \frac{0.027}{1.000}$

$$f(x) = \frac{n!}{x!(n-x)!} p^x (1-p)^{(n-x)}$$

If you Send n = 1000 Emails

X = 301 responses

$$f(301) = \binom{1000}{301} (.3)^{301} (.7)^{1000-301}$$

$$= \frac{1000!}{301!\,(1000-301)!} (.3)^{301} (.7)\text{^}699$$

$$= .027$$

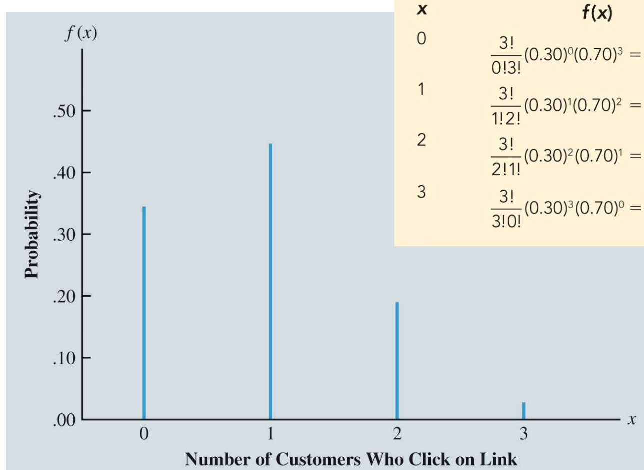
Or 2.7% Chance

Discrete Probability Distributions

Probability Distribution for the Number of Customers Who Click on the Link in the Martin's Targeted E-Mail

(3 Emails Sent)

- What's the probability that NO MORE THAN 1 person response?
  - 0.343 + 0.441 = 0.784
  - Or 78.4% chance that 1 or none respond
  - (Add the probabilities together)



x	f(x)
0	$\frac{3!}{0!3!}(0.30)^0(0.70)^3 = 0.343$
1	$\frac{3!}{1!2!}(0.30)^1(0.70)^2 = 0.441$
2	$\frac{3!}{2!1!}(0.30)^2(0.70)^1 = 0.189$
3	$\frac{3!}{3!0!}(0.30)^3(0.70)^0 = \frac{0.027}{1.000}$

Discrete Probability Distributions

Figure 4.15: Excel Worksheet for Computing Binomial Probabilities of the Number of Customers Who Make a Purchase at Martin's

	A	B	C	D
1			Number of Trials (n):	3
2			Probability of Success (p):	0.3
3				
4	x	$f(x)$	Cumulative Probability	
5	0	=BINOM.DIST(A5,$D$1,$D$2,FALSE)	=BINOM.DIST(A5,$D$1,$D$2,TRUE)	
6	1	=BINOM.DIST(A6,$D$1,$D$2,FALSE)	=BINOM.DIST(A6,$D$1,$D$2,TRUE)	
7	2	=BINOM.DIST(A7,$D$1,$D$2,FALSE)	=BINOM.DIST(A7,$D$1,$D$2,TRUE)	
8	3	=BINOM.DIST(A8,$D$1,$D$2,FALSE)	=BINOM.DIST(A8,$D$1,$D$2,TRUE)	

	A	B	C	D
1			Number of Trials (n):	3
2			Probability of Success (p):	0.3
3				
4	x	$f(x)$	Cumulative Probability	
5	0	0.343	0.343	
6	1	0.441	0.784	
7	2	0.189	0.973	
8	3	0.027	1.000	

Discrete Probability Distributions

Poisson Probability Distribution:

- useful in estimating the number of occurrences of an event over a specified interval of time and space.
- Examples:
  - Number of patients who arrive at a health care clinic between 8:00am and 9:00am.
  - Number of computer-server failures in a month.
  - Number of repairs needed in 10 miles of highway.
  - Number of leaks in 100 miles of pipeline.

Discrete Probability Distributions

- **Poisson probability distribution applies if:**
  - 1. The probability of an occurrence is the same for any two intervals (of time or space) of equal length.
    - Probability of a patient arriving between 8:00-8:01 = probability of a patient arriving between 8:59-9:00 am
  - 2. The occurrence or nonoccurrence in any interval (of time or space) is independent of the occurrence or nonoccurrence in any other interval.
    - The arrival of a patient at 8:02am is independent of a patient arriving at 8:03am

POISSON PROBABILITY MASS FUNCTION

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} \quad \textbf{(4.17)}$$

where

$f(x)$ = the probability of x occurrences in an interval
μ = expected value or mean number of occurrences in an interval
$e \approx 2.71828$

Discrete Probability Distributions

- **Poisson probability example:**

Average number of patients arriving in a 15 minute period = 10.

POISSON PROBABILITY MASS FUNCTION

$$f(x) = \frac{\mu^x e^{-\mu}}{x!} \quad \textbf{(4.17)}$$

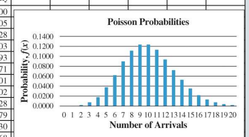
where

$f(x)$ = the probability of x occurrences in an interval
μ = expected value or mean number of occurrences in an interval
$e \approx 2.71828$

$$\text{Probability of exactly5arrivals in15 minutes} = f(5) = \frac{10^5 e^{-10}}{5!} = 0.0378$$

Discrete Probability Distributions

Figure 4.16: Excel Worksheet for Computing Poisson Probabilities of the Number of Patients Arriving at the Emergency Room



	A	B	C	D
1			Mean Number of Occurrences:	10
2				
3	Number of Arrivals	Probability, $f(x)$	Cumulative Probability	
4	0	=POISSON.DIST(A4,$D$1,FALSE)	=POISSON.DIST(A4,$D$1,TRUE)	
5	1	=POISSON.DIST(A5,$D$1,FALSE)	=POISSON.DIST(A5,$D$1,TRUE)	
6	2	=POISSON.DIST(A6,$D$1,FALSE)	=POISSON.DIST(A6,$D$1,TRUE)	
7	3	=POISSON.DIST(A7,$D$1,FALSE)	=POISSON.DIST(A7,$D$1,TRUE)	
8	4	=POISSON.DIST(A8,$D$1,FALSE)	=POISSON.DIST(A8,$D$1,TRUE)	
9	5	=POISSON.DIST(A9,$D$1,FALSE)	=POISSON.DIST(A9,$D$1,TRUE)	
10	6	=POISSON.DIST(A10,$D$1,FALSE)	=POISSON.DIST(A10,$D$1,TRUE)	
11	7	=POISSON.DIST(A11,$D$1,FALSE)	=POISSON.DIST(A11,$D$1,TRUE)	
12	8	=POISSON.DIST(A12,$D$1,FALSE)	=POISSON.DIST(A12,$D$1,TRUE)	
13	9	=POISSON.DIST(A13,$D$1,FALSE)	=POISSON.DIST(A13,$D$1,TRUE)	
14	10	=POISSON.DIST(A14,$D$1,FALSE)	=POISSON.DIST(A14,$D$1,TRUE)	
15	11	=POISSON.DIST(A15,$D$1,FALSE)	=POISSON.DIST(A15,$D$1,TRUE)	
16	12	=POISSON.DIST(A16,$D$1,FALSE)	=POISSON.DIST(A16,$D$1,TRUE)	
17	13	=POISSON.DIST(A17,$D$1,FALSE)	=POISSON.DIST(A17,$D$1,TRUE)	
18	14	=POISSON.DIST(A18,$D$1,FALSE)	=POISSON.DIST(A18,$D$1,TRUE)	
19	15	=POISSON.DIST(A19,$D$1,FALSE)	=POISSON.DIST(A19,$D$1,TRUE)	
20	16	=POISSON.DIST(A20,$D$1,FALSE)	=POISSON.DIST(A20,$D$1,TRUE)	
21	17	=POISSON.DIST(A21,$D$1,FALSE)	=POISSON.DIST(A21,$D$1,TRUE)	
22	18	=POISSON.DIST(A22,$D$1,FALSE)	=POISSON.DIST(A22,$D$1,TRUE)	
23	19	=POISSON.DIST(A23,$D$1,FALSE)	=POISSON.DIST(A23,$D$1,TRUE)	
24	20	=POISSON.DIST(A24,$D$1,FALSE)	=POISSON.DIST(A24,$D$1,TRUE)	

	A	B	C	D
1			Mean Number of Occurrences:	10
2				
3	Number of Arrivals	Probability, $f(x)$	Cumulative Probability	
4	0	0.0000	0.0000	
5	1	0.0005	0.0005	
6	2	0.0023	0.0028	
7	3	0.0076	0.0103	
8	4	0.0189	0.0293	
9	5	0.0378	0.0671	
10	6	0.0631	0.1301	
11	7	0.0901	0.2202	
12	8	0.1126	0.3328	
13	9	0.1251	0.4579	
14	10	0.1251	0.5830	
15	11	0.1137	0.6968	
16	12	0.0948	0.7916	
17	13	0.0729	0.8645	
18	14	0.0521	0.9165	
19	15	0.0347	0.9513	
20	16	0.0217	0.9730	
21	17	0.0128	0.9857	
22	18	0.0071	0.9928	
23	19	0.0037	0.9965	
24	20	0.0019	0.9984	

Continuous Probability Distributions

Uniform Probability Distribution

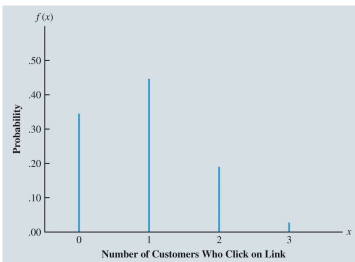
Triangular Probability Distribution

Normal Probability Distribution

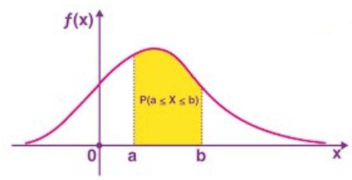
Exponential Probability Distribution

Continuous Probability Distributions

- Probability of a **Discrete** Random Variable
  - the **probability mass function f(x)** provides the probability that the random variable assumes a particular value.



- Probability of a **Continuous** Random Variable
  - the **probability density function**, **f(x)** does not directly provide probabilities
    - the probability of any particular value of the random variable is zero
    - We are computing the probability that the random variable assumes any value in an interval.



Continuous Probability Distributions

Continuous Uniform Probability Distribution:

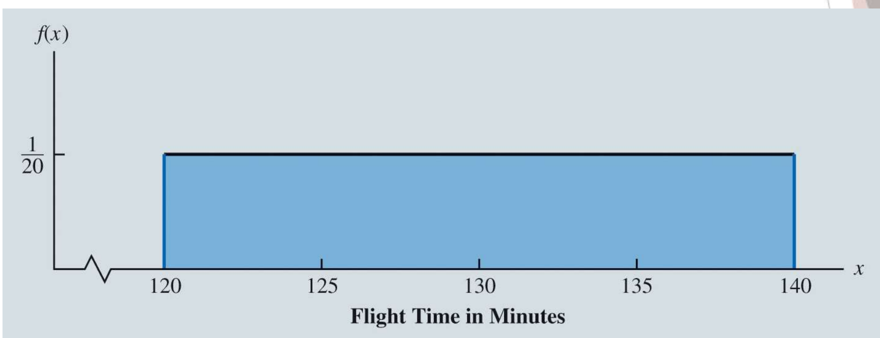
- Example:
  - Random variable *x* representing the flight time of an airplane traveling from Chicago to New York.
  - Suppose the flight time can be any value between 120 and 140 minutes
  - Assuming the probability of any interval inside of the 120 – 140 minutes is the same
    - This implies => Uniform Distribution

UNIFORM PROBABILITY DENSITY FUNCTION

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for } a \le x \le b \\ 0 & \text{elsewhere} \end{cases} \qquad \textbf{(4.18)}$$

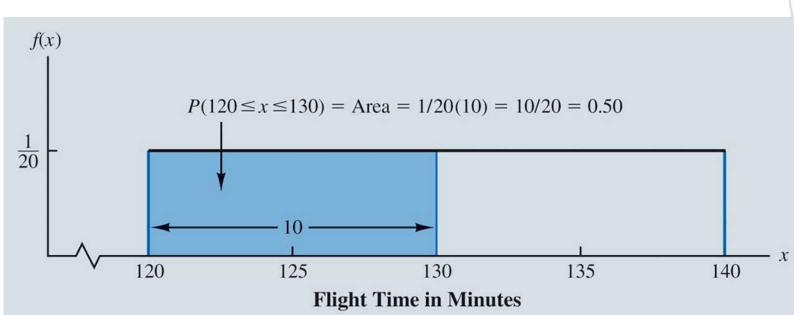
Continuous Probability Distributions

Figure 4.17: Uniform Probability Distribution for Flight Time



Continuous Probability Distributions

Figure 4.18: The Area Under the Graph Provides the Probability of a Flight Time Between 120 and 130 Minutes



Continuous Probability Distributions

UNIFORM DISTRIBUTION: CUMULATIVE PROBABILITIES

$$P(x \leq x_0) = \frac{x_0 - a}{b - a} \quad \text{for } a \leq x_0 \leq b \qquad \textbf{(4.19)}$$

- The calculation of the expected value and variance for a continuous random variable is analogous to that for a discrete random variable.
- For uniform continuous probability distribution, the formulas for the expected value and variance are:

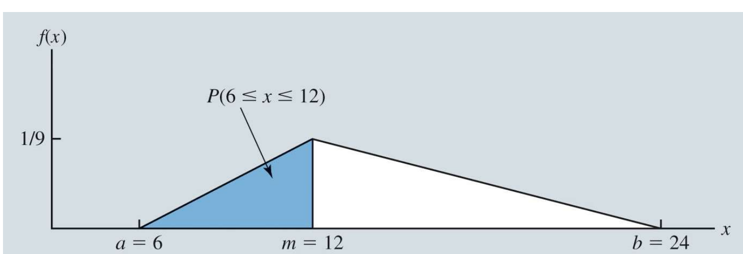
$$E(x) = \frac{a+b}{2} = \frac{120 + 140}{2} = 130$$

$$\text{Var}(x) = \frac{(b-a)^2}{12} = \frac{(140 - 120)^2}{12} = 33.33$$

Continuous Probability Distributions

Triangular Probability Distribution:

- Useful only when subjective probability estimates are available.
- In the **triangular probability distribution**, we need only specify:
  - The minimum possible value *a*.
  - The maximum possible value *b*.
  - The most likely value (or mode) of the distribution *m*.



Continuous Probability Distributions

- The general form of the triangular probability density function is:

TRIANGULAR PROBABILITY DENSITY FUNCTION

$$f(x) = \begin{cases} \dfrac{2(x-a)}{(b-a)(m-a)} & \text{for } a \le x \le m \\ \dfrac{2(b-x)}{(b-a)(b-m)} & \text{for } m < x \le b \end{cases} \qquad \textbf{(4.20)}$$

where

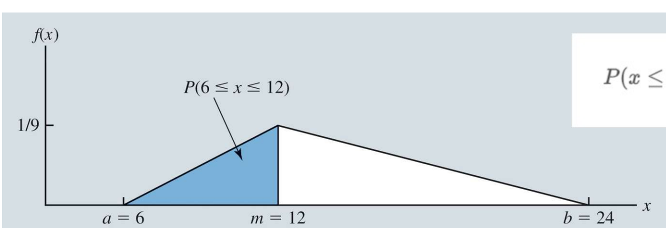
a = minimum value
b = maximum value
m = mode

Continuous Probability Distributions

- The geometry required to find the area under the graph for any given value is slightly more complex than that required to find the area for a uniform distribution.

TRIANGULAR DISTRIBUTION: CUMULATIVE PROBABILITIES

$$P(x \le x_0) = \begin{cases} \dfrac{(x_0-a)^2}{(b-a)(m-a)} & \text{for } a \le x_0 \le m \\ 1 - \dfrac{(b-x_0)^2}{(b-a)(b-m)} & \text{for } m < x_0 \le b \end{cases} \qquad \textbf{(4.21)}$$

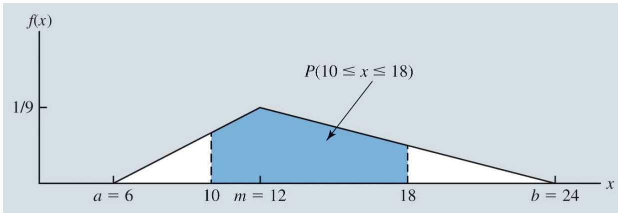


$$P(x \le 12) = \frac{(12-6)^2}{(24-6)(12-6)} = 0.3333$$

Continuous Probability Distributions

Figure 4.20: Triangular Distribution to Determine

$$P(10 \le x \le 18) = P(x \le 18) - P(x \le 10)$$

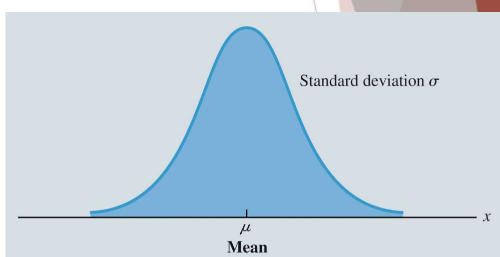


$$P(x \le 18) - P(x \le 10) = \left[1 - \frac{(24-18)^2}{(24-6)(24-12)}\right] - \left[\frac{(10-6)^2}{(24-6)(10-6)}\right] = 0.6111$$

Continuous Probability Distributions

Normal Probability Distribution:

- One of the most useful probability distributions
- Wide variety of practical and business applications:
  - Heights and weights of people.
  - Test scores.
  - Scientific measurements.
  - Uncertain quantities such as demand for products.
  - Rate of return for stocks and bonds.
  - Time it takes to manufacture a part or complete an activity.



Continuous Probability Distributions

- The probability density function that defines the bell-shaped curve of the normal distribution is:

NORMAL PROBABILITY DENSITY FUNCTION

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-(x-\mu)^2/2\sigma^2} \tag{4.22}$$

where

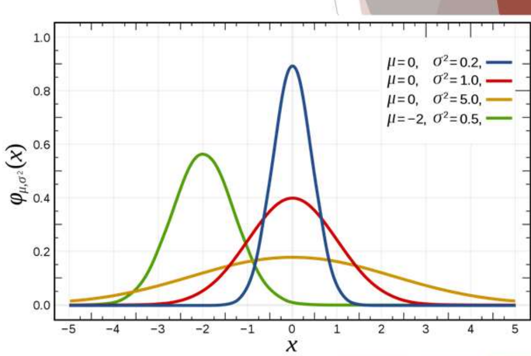
μ = mean
σ = standard deviation
$\pi \approx 3.14159$
$e \approx 2.71828$

Continuous Probability Distributions

Characteristics of the normal distribution:

1. The entire family of normal distributions is differentiated by two parameters

$$\mu \quad \text{and} \quad \sigma$$



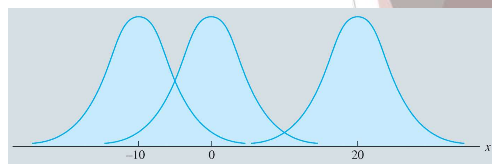
2. The highest point on the normal curve is at the mean, which is also the median and mode of the distribution.
3. The mean of the distribution can be any numerical value: negative, zero, or positive

Continuous Probability Distributions

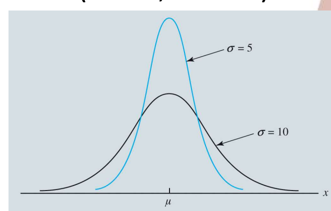
Characteristics of the normal distribution (continued):

4. The normal distribution is symmetric
5. The tails of the curve extend to infinity in both directions and theoretically never touch the horizontal axis.
6. The standard deviation determines how flat and wide the normal curve is;
   - larger values of the standard deviation result in wider, flatter curves
   - showing more variability in the data

$(\mu = -10,\ \mu = 0,\ \mu = 20)$



$(\sigma = 5,\ \sigma = 10)$



Continuous Probability Distributions

Characteristics of the normal distribution (continued):

7. Probabilities = areas under the normal curve.
   - The total area under the curve for the normal distribution is 1.
8. The percentages of values in some commonly used intervals are:

   a. 68.3% of the values are within

   + or – 1 SD of the mean

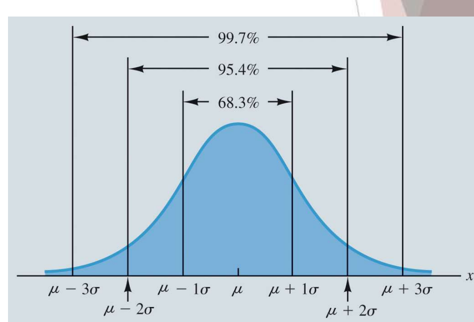
   b. 95.4% of the values are within

   + or – 2 SD of the mean

   C. 99.7% of the values are within

   + or – 3 SD of the mean

Areas Under the Curve for
Any Normal Distribution



Continuous Probability Distributions

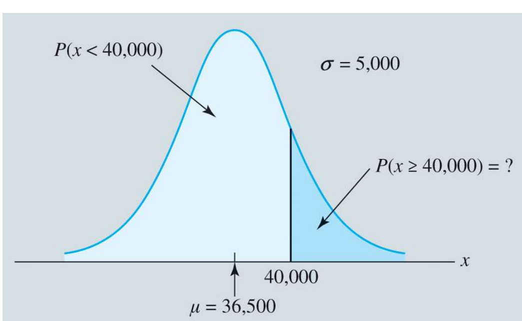
- Example: Grear Aircraft Engines sells aircraft engines to commercial airlines.
  - Grear offers guarantees that engines will provide certain amount of lifetime flight hours
  - Based on extensive flight testing and computer simulations, Grear believes mean lifetime flight hours is normally distributed with a mean

$\mu = 36{,}500$ hours and standard deviation $\sigma = 5{,}000$ hours.

- What is the probability that an engine will last more than 40,000 hours?

Continuous Probability Distributions

Figure 4.25: Grear Aircraft Engines Lifetime Flight Hours Distribution



Calculate P(x >= 40,000) by hand:

- Standardize 40,000 as a Z-score
- Look up the z-score on the Z-table
- Find its corresponding probability
  - Row corresponding to leading digit
  - Column corresponding to second digit
  - Intersect row and column
- For P(x>=40,000) = 1 – (intersection of row and column)

Continuous Probability Distributions

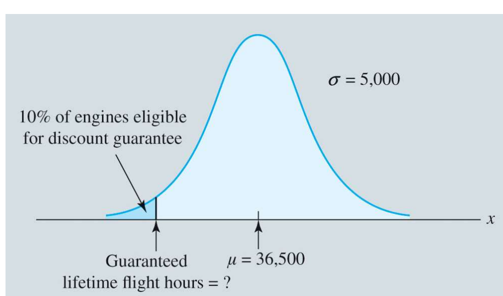
Figure 4.26: Excel Calculations for Grear Aircraft Engines Example

	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	**$P(x \leq 40{,}000)$ =**	=NORM.DIST(40000, \$B\$1, \$B\$2,TRUE)	
6	**$P(x > 40{,}000) = 1 - P(x \leq 40{,}000)$ =**	=1-B5	
7			
8	**Guarantee on Lifetime Flight Hours for 10% of engines eligible for discount guarantee:**	=NORM.INV(0.1, \$B\$1, \$B\$2)	

	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	**$P(x \leq 40{,}000)$ =**	0.7580	
6	**$P(x > 40{,}000) = 1 - P(x \leq 40{,}000)$ =**	0.2420	
7			
8	**Guarantee on Lifetime Flight Hours for 10% of engines eligible for discount guarantee:**	30092.24	

Continuous Probability Distributions

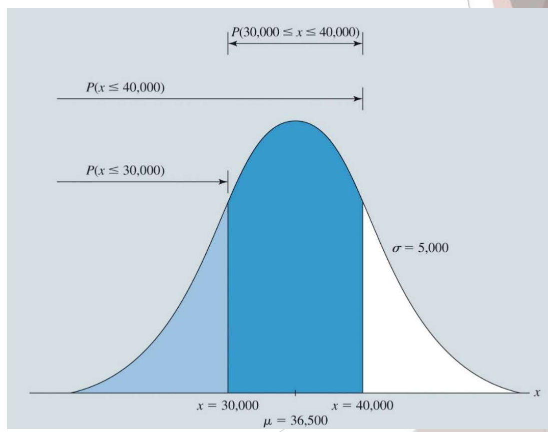
- Grear is considering a guarantee that will provide a discount on a replacement aircraft engine if the original engine does not meet the lifetime-flight-hour guarantee.
- How many lifetime flight hours should Grear guarantee if Grear wants no more than 10% of aircraft engines to be eligible for the discount guarantee?



Continuous Probability Distributions

How do we calculate the probability that an engine will have a lifetime of flight hours greater than 30,000 but less than 40,000 hours?

$$P(30{,}000 \le x \le 40{,}000)$$



Continuous Probability Distributions

Using Excel to Find Aircraft Engines Example

	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	***P* (*x* ≤ 40,000) =**	=NORM.DIST(40000, $B$1, $B$2,TRUE)	
6	***P* (*x* ≤ 30,000) =**	=NORM.DIST(30000, $B$1, $B$2,TRUE)	
7			
8	***P* (30,000 ≤ *x* ≤ 40,000) = *P* (*x* ≤ 40,000) − *P* (*x* ≤ 30,000) =**	=B5-B6	

	A	B	C
1	**Mean:**	36500	
2	**Standard Deviation:**	5000	
3			
4			
5	***P* (*x* ≤ 40,000) =**	0.7580	
6	***P* (*x* ≤ 30,000) =**	0.0968	
7			
8	***P* (30,000 ≤ *x* ≤ 40,000) = *P* (*x* ≤ 40,000) − *P* (*x* ≤ 30,000) =**	0.6612	

Continuous Probability Distributions

Exponential Probability Distribution:

- may be used for random variables such as:
  - Time between patient arrivals at an emergency room.
  - Distance between major defects in a highway.
  - Time until default in certain credit-risk models.

EXPONENTIAL PROBABILITY DENSITY FUNCTION

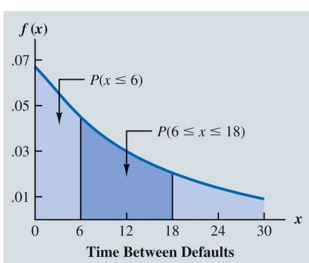
$$f(x) = \frac{1}{\mu} e^{-x/\mu} \quad \text{for } x \geq 0 \tag{4.23}$$

where

μ = expected value or mean
e = 2.71828

Continuous Probability Distributions

- Example: *x* represents time between business loan defaults for a particular lending agency.
- If the mean time between loan defaults is 15 months,



What is the probability that the next default will be between 6 and 18 months?

- To compute exponential probabilities, we use:

EXPONENTIAL DISTRIBUTION: CUMULATIVE PROBABILITIES

$$P(x \leq x_0) = 1 - e^{-x_0/\mu} \tag{4.24}$$

Continuous Probability Distributions

Figure 4.31: Using Excel to Calculate Business Loan Defaults Example

	A	B	C
1	**Mean, μ =**	15	
2			
3	P ($x \le 18$) =	=EXPON.DIST(18,1/$B$1, TRUE)	
4	P ($x \le 6$) =	=EXPON.DIST(6,1/$B$1, TRUE)	
5	P ($6 \le x \le 18$) = P ($x \le 18$) – P ($x \le 6$) =	=B3-B4	

	A	B	C
1	**Mean, μ =**	15	
2			
3	P ($x \le 18$) =	0.6988	
4	P ($x \le 6$) =	0.3297	
5	P ($6 \le x \le 18$) = P ($x \le 18$) – P ($x \le 6$) =	0.3691	

- Figure 4.31 shows how to calculate these values for an exponential distribution in Excel using *EXPON.DIST.*