Chapter Five

Discrete and Continuous Probability Distributions

Chapter Goals:

After completing this chapter, you should be able to

- Apply the binomial distribution to applied problems.
- Compute probabilities for The Poisson and Hypergeometric distributions.
- Find the probabilities using a normal distribution table and the normal business problems.
- Recognize when to apply the uniform and exponential distributions.

Probability Distributions:

1. Discrete Probability Distributions:

1. Binomial Distribution.
2. Poisson Distribution.
3. Hypergeometric Distribution.

Discrete Random Variable: It is a random variable that can assume only *countable* number of possible values.

Examples: 1. Number of students who come late to the class.
2. Number of T.V sets in a household.
3. Number of customers who arrive at a store.

In general: There are only TWO possible outcomes.

Gender: Male or Female.

Defective: Yes or No.

2. Continuous Probability Distributions:

1. Normal Distribution.
2. Uniform Distribution.
3. Exponential Distribution.

Continuous Random Variable: it is a random variable that can assume any value in *an interval*. (**Uncountable number of possible values**)

Examples: 1. Height and length.
2. Time required to complete a job.
3. Temperature.

In general: These can potentially take on any value, depending only on the ability to measure accurately.

5.1: The Binomial Probability Distribution:

Characteristics of the Binomial Distribution:

1. A trial has only two possible outcomes "***S**uccess*" or "***F**ailure*".
2. There is fixed number of trials "*n*".
3. The trials of the experiment are independent of each other (sampling is conducted with replacement).
4. The probability of "***S***" for each trial is *p* and it is fixed, and the probability of "***F***" is $q = 1 - p$.

Binomial Distribution Formula:

$$P(x) = C_x^n p^x (1-p)^{n-x} \quad , \; x = 0,1,2,\ldots n$$

Where:

$C_x^n = \dfrac{n!}{x!(n-x)!}$ Number of combinations of x objects from n objects.

$n! = n(n-1)(n-2)\ldots(2)(1)$, $x! = x(x-1)(x-2)\ldots(2)(1)$, $0! = 1$

and

$P(x)$: Probability of getting x successes in n trials.

x: Number of successes in the experiment.

n: Number of trials (sample size).

p: Probability of success for *one* trial.

Mean and Variance for the Binomial Distribution:

Mean (Expected value or average): $\mu = np$.

Variance: $\sigma^2 = np(1-p)$ and standard deviation $\sigma = \sqrt{np(1-p)}$.

Example: Tossing a *fair* coin 5 times, let *X*: number of H's comes up;

1. Write the sample space

2. Find probability to get 3 H's exact.

3. Find the probability at most 5 H's exist.

4. Write the probability distribution, and identify it, for the r.v. *X*.

5. Find the mean and the standard deviation for *X*.

Example: An Urn contains 10 calculators, 4 of the calculators are defective. If a sample, of size 4, is randomly selected *with replacement*, Find:

1. The probability that two of them are defective.

2. The probability that at least two are defective.

3. The mean and the standard deviation for the number of defective calculators.

4. The coefficient of variation for the number of defective calculators.

Example: Oil drilling company ventures into various locations, and their success or failure is independent from one location to another. Suppose the probability of hitting a productive well at any specific location is 0.1. If a driller drills 5 locations,

1. Find the probability that there will be at least two productive wells.

2. Find the probability that between two and four, inclusive, productive wells will be hit.

3. Complete the probabilities for the rest of the number of productive wells.

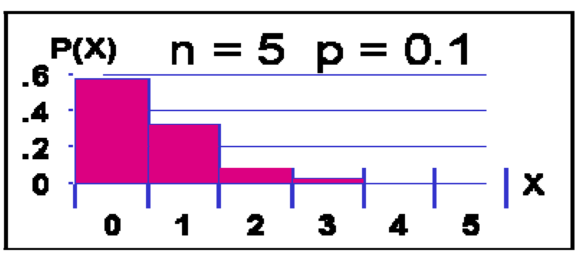
Example: Q 5.13 Page 187

Example: Q5.22 Page 188

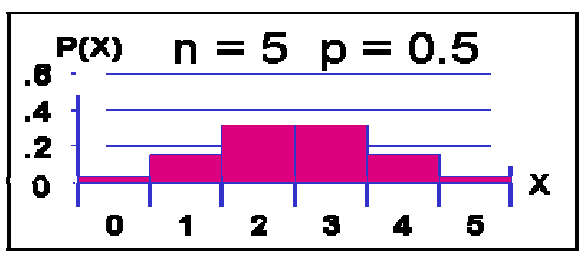
Note:

The shape of the binomial distribution depends on the value of *p*.

1. If $p = 0.5$ then the shape is *symmetric* regardless of the sample size.
2. If $p < 0.5$ then the shape is *right-skewed* regardless of the sample size.
3. If $p > 0.5$ then the shape is *left-skewed* regardless of the sample size.



Here $n = 5$ and $p = 0.1$



Here $n = 5$ and $p = 0.5$

5.2: The Poisson Probability Distribution:

Characteristics of the Poisson Distribution:

1. The outcomes of interest are relative to the possible outcomes.
2. The average number of outcomes of interest per time or some *space interval* is λ .
3. The number of outcomes of interest is random, and the su.

4. The probability of an outcome of interest occurs in a given segment is the same for all segments.

Poisson Distribution Formula:

$$P(x) = \frac{(\lambda t)^x e^{-\lambda t}}{x!}, \quad x = 0, 1, 2, \ldots$$

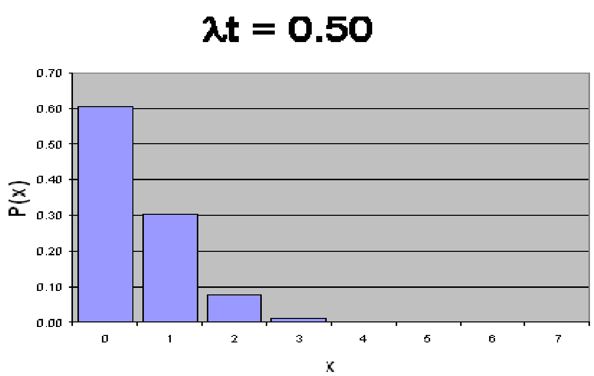
Where:
x: Number of events in a segment (interval) of a space.
t: Size of the segment of interest (number of segments).
λ: Expected (average) number of events in a segment of unit size.
e: Base of the natural logarithmic function.

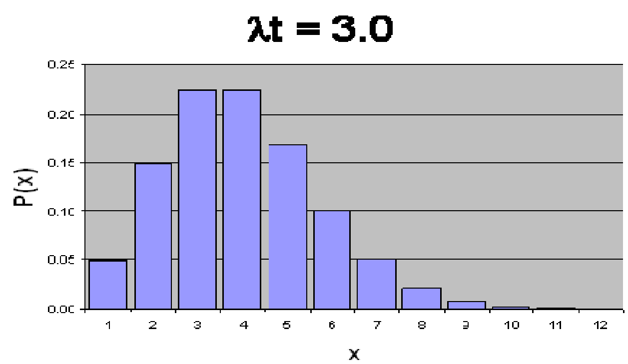
Mean and Variance of the Poisson Distribution:

Mean: $\mu = \lambda t$.

Variance: $\sigma^2 = \lambda t$ and standard deviation $\sigma = \sqrt{\lambda t}$.

Note: The shape of the Poisson distribution depends (?) on the two parameters λ and t, i.e. the *mode* will always equal to λt:





Example: At a checkout center, customers arrive at an average of 2.5 *customers* per minute.

1. What is the probability that two will arrive within 2 minutes?

2. What is the probability that at least two will arrive within 3 minutes?

3. Find the standard deviation of the number of customers who will arrive to the counter within one and a half hours.

Example: The number of failures of a testing instrument from contamination particles on the product is a Poisson random variable with a mean of 0.02 failures per hour.

a. What is the probability that the instrument does not fail in an 8-hour shift?

Let X = number of failures of the testing instrument per hour. Then the expected number of failures in an 8-hour shift is given by
$\lambda t = 0.02(8) = 0.16$, so that $P(X = 0) = e^{-\lambda} = e^{-0.16} = 0.852$.

b. What is the probability of at least one failure in 30 minutes?

The expected number of failures in 30 minutes is given by $\lambda t = 0.02(30/60) = 0.01$ so that $P(X \geq 1) = 1 - P(X = 0) = 1 - e^{-\lambda} = 1 - e^{-0.01} = 0.01$.

Example: On the average, a certain intersection results in 3 traffic accidents per month. What is the probability that for any given month at this intersection;

a. Exactly five accidents will occur in a month?

b. Less than three accidents will occur in two months?

c. At least three accidents will occur in three months?

The Hypergeometric Distribution

Characteristics of the Hypergeometric Distribution:

- n trials in a sample are taken from a *finite population* of size N.
- Sample is taken *without replacement*.
- Therefore the trials are *dependent*.
- Concerned with finding the probability of x successes in the sample from X successes in the population.

Hypergeometric Distribution Formula

$$P(x) = \frac{C_{n-x}^{N-X} \; C_x^X}{C_n^N}$$ (Two possible outcomes per trial)

Where

N = Population size.

X = number of successes in the population.

n = sample size.

x = number of successes in the sample.

$n - x$ = number of failures in the sample.

Example: Three Light bulbs were selected, *without replacement*, from a box of ten. Of the ten there were 4 defectives. What is the probability that two, of the three selected, were defective?

Solution:

$N = 10$ $\qquad$ $n = 3$

$X = 4$ $\qquad$ $x = 2$

$$P(x=2) = \frac{C_{n-x}^{N-X} \; C_x^X}{C_n^N} = \frac{C_1^6 \; C_2^4}{C_3^{10}} = \frac{(6)(6)}{120} = 0.3$$

Example: An Urn contains ten calculators, three of them are defective. If a sample of size four is randomly selected, *without replacement*, find:

a. The probability that 2 of them defective

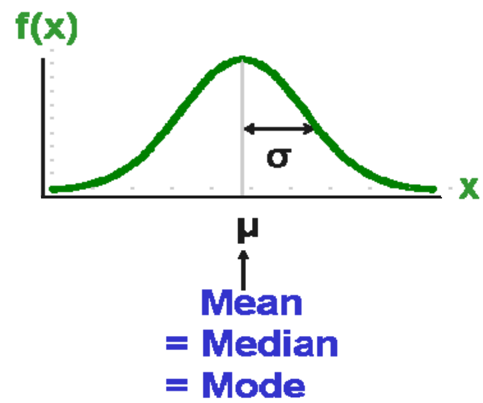
b. The probability that at least 2 defective.

c. The probability that 2 or 3 defective

Example: Q 5.33 Page 198

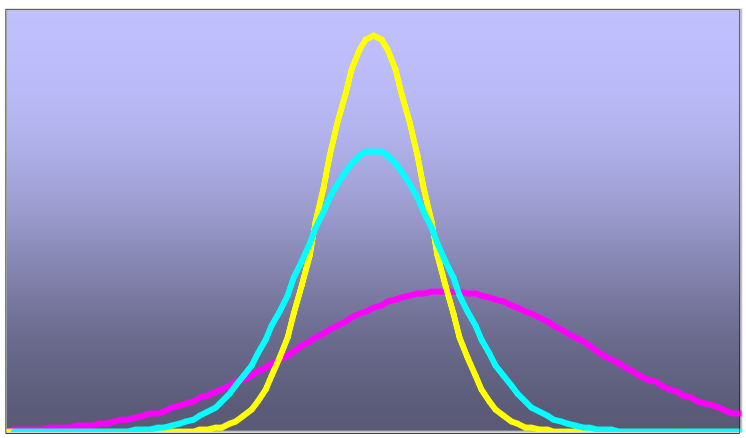
5.3: The Normal Distribution:

- Bell Shaped (symmetric and uni-modal).
- Mean, Median and Mode are equal.
- Location is determined by the mean, μ.
- Spread (or variability) is determined by the standard deviation, σ.
- The random variable has an infinite theoretical domain: $(-\infty, +\infty)$.

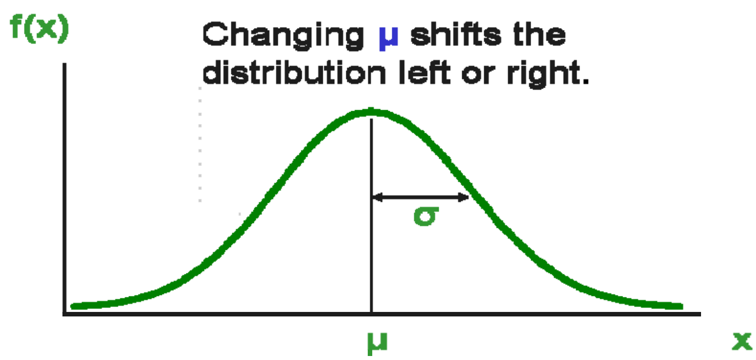


Note:

By varying the parameters μ and σ, we obtain different normal distributions

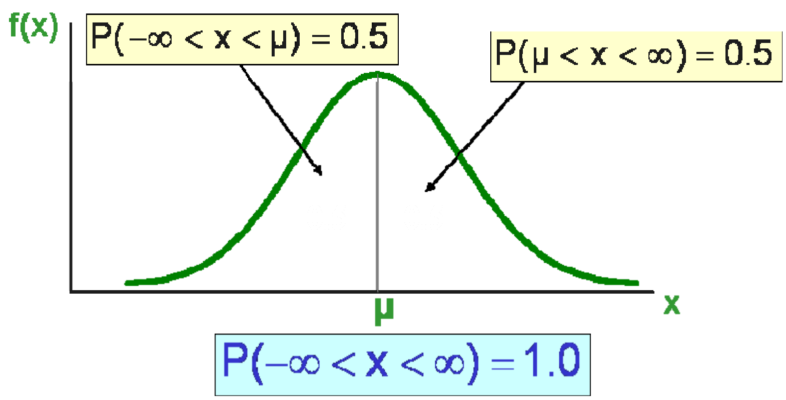


The Normal Distribution Shape



Finding Normal Probabilities

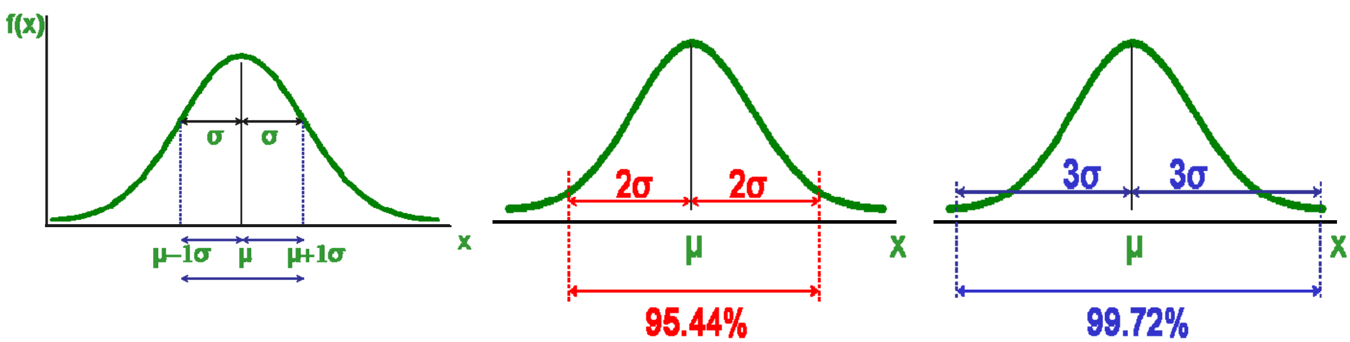
- Probability is measured by the *area* under the curve.
- The total area under the curve is 1.0, and the curve is symmetric about the mean, so half is above the mean, half is below.



Empirical Rules

What can we say about the distribution of values around the mean? There are some general rules:

1. $[\mu \pm 1\sigma]$ encloses about 68% of x's.
2. $[\mu \pm 2\sigma]$ covers about 95% of x's.
3. $[\mu \pm 3\sigma]$ covers about 99.7% of x's.

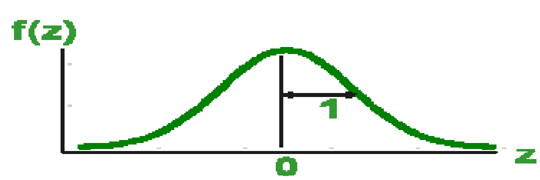


- If a value is about 2 or more standard deviations away from the mean in a normal distribution, then it is considered far from the mean.

- The chance that a value is far or farther away from the mean is highly unlikely, given a particular mean and standard deviation.

The Standard Normal Distribution

- Also known as the Z distribution.
- Mean is defined to be 0.
- Standard Deviation is 1.



Note:

Values above the mean have positive z-values, values below the mean have negative z-values.

Example: Let Z has a standard Normal distribution, find the following:

a. $P(0 < Z < 1.34) =$

b. $P(Z > 2.56) =$

c. $P(Z < 1.78) =$

d. $P(1.45 < Z < 3.04) =$

e. $P(-0.93 < Z < 0) =$

f. $P(Z < -1.33) =$

g. $P(Z > -3.05) =$

h. $P(-2.45 > Z > -3.05) =$

i. $P(-1.45 < Z < 3.04) =$

Translation to the Standard Normal Distribution

- Any normal distribution (with any mean and standard deviation combination) can be transformed into the standard normal distribution (Z).
- Need to transform X units into Z units.
- Translate from X to the standard normal (the Z distribution) by subtracting the mean of x and dividing by its standard deviation:

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

Example: If X is distributed normally with mean of 100 and standard deviation 5, find the following:

a. $P(X > 102) =$

b. $P(X > 96) =$

c. $P(85 < X < 104) =$

d. The median

e. The 8th decile.

f. Find a such that $P(X > a) = 0.05$

g. If a sample of size 10 is randomly selected, find the probability that 3 have a value of at least 102.

5.4: The Uniform Distribution:
The uniform distribution is a probability distribution that has *equal probabilities* for all possible outcomes of the random variable.

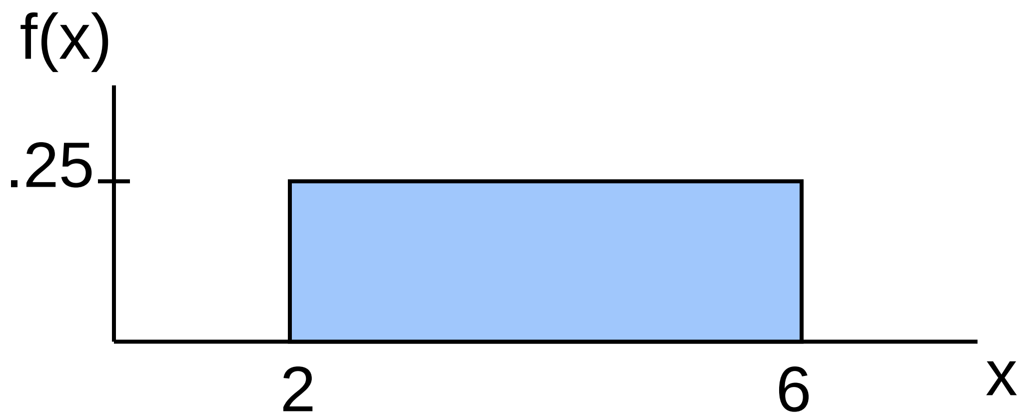
The Continuous Uniform Distribution

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

where

$f(x)$ = value of the density function at any x value.
a = lower limit of the interval.
b = upper limit of the interval.

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



The mean and the standard deviation for the Uniform distribution:

Expected value (mean): $E(X) = \mu_X = \dfrac{a+b}{2}$

Standard deviation: $V(X) = \sigma_X = \sqrt{\dfrac{(b-a)^2}{12}}$

Example: Q5.69 Page 216:

a. $P(X > 50) = (60 - 50)/(60 - 20) = 0.25$

b. $P(X = 45) = 0$; you cannot find the probability of a specific value in a continuous distribution.

c. $P(25 < X < 35) = (35 - 25)/(60 - 20) = 0.25$

d. $P(X < 34) = (34 - 20)/(60 - 20) = 0.35$

The Exponential Distribution:
Used to measure the space length that elapses between two occurrences of an event (the time between arrivals).

Examples:

- Time between trucks arriving at an unloading dock.
- Time between transactions at an ATM Machine.
- Time between phone calls to the main operator.

The Exponential Distribution:
A continuous r.v. that has exponential distribution (X: Exp (λ)) has the probability *density* function (pdf) given by:

$$f(x) = \lambda e^{-\lambda x}, \quad x \geq 0.$$

Where λ: The average number of events in a space segment.

Where 1. Mean: $\mu = \frac{1}{\lambda}$.

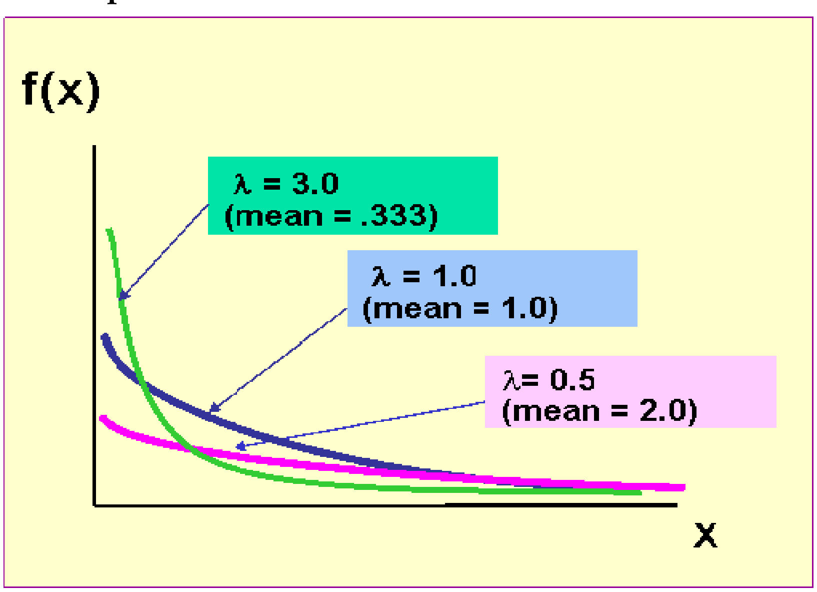
2. Variance: $\sigma^2 = \frac{1}{\lambda^2}$, and the standard deviation $\sigma = \frac{1}{\lambda}$.

Notes:

1. The probability that an arrival time between two specified time a is:
$P(a \leq X \leq b) = \int_a^b \lambda e^{-\lambda x} dx = -e^{-\lambda b} - \left(-e^{-\lambda a}\right) = e^{-\lambda a} - e^{-\lambda b}$.
2. The probability that an arrival time is equal to or less than some specified time a is: $P(X < b) = P(0 \leq X \leq b) = e^{-\lambda 0} - e^{-\lambda b} = 1 - e^{-\lambda b}$.
3. The probability that an arrival time is grater than some specified time a is:
$P(X > a) = P(a < X < \infty) = e^{-\lambda a} - e^{-\lambda \infty} = e^{-\lambda a} - 0 = e^{-\lambda a}$.

Note: If the number of occurrences per space segment is a Poisson with mean λ, then the space length between occurrences is exponential with mean space length equals to $1/\lambda$.

Shape of the exponential distribution



Example: Q5.70 Page217

a. $P(X > 2) = e^{-(1/2)(2)} = 0.3678$

b. $P(1<X<2) = P(X<2) - P(X<1) = 0.6321 - 0.3935 = 0.2386$

c. $P(X>2.5) = e^{-(1/2)(2.5)} = 0.2865$

d. Find the median

$$P(X < a) = 0.5$$
$$1 - e^{-2a} = 0.5 \Rightarrow e^{-2a} = 0.5 \Rightarrow -2a = \ln(0.5) = -0.69314718 \Rightarrow a = 0.34657$$

Example: Customers arrive at the claims counter at the rate of 15 per hour (Poisson process). What is the probability that the arrival time between consecutive customers is less than five minutes?

- Time between arrivals is exponentially distributed with mean time between arrivals of 4 minutes (15 per 60 minutes, on average)
- $1/\lambda = 4.0$, so $\lambda = .25$
- $P(X < 5) = 1 - e^{-\lambda a} = 1 - e^{-(.25)(5)} = .7135$

Example: Q5.74 Page 217

Poisson Distribution with $\lambda = 60$ per arrivals per 60 minutes = $\lambda = 1.0$ per minute.

a. $P(X > 1) = 1 - P(X \leq 1)$

$P(X \leq 1) = 1 - .3679 = 0.6321$

$P(X > 1) = 1 - 0.6321 = 0.3679$

b. 45 seconds = .75 minutes and 75 seconds = 1.25 minutes

$P(.75 \leq X \leq 1.25) = ?$

$P(X \leq 1.25) = 1 - 0.2865 = 0.7135$

$P(X < .75) = 1 - 0.4724 = 0.5276$

$P(.75 \leq X \leq 1.25) = 0.7135 - 0.5276 = 0.1859$