Probability Distribution

What is Distribution

- Describes the ‘shape’ of a batch of numbers
- The characteristics of a distribution can sometimes be defined using a small number of numeric descriptors called ‘parameters’

Why is Distribution?

- Can serve as a basis for standardized comparison of empirical distributions
- Can help us estimate confidence intervals for inferential statistics
- Form a basis for more advanced statistical methods
  - ‘fit’ between observed distributions and certain theoretical distributions is an assumption of many statistical procedures

Random Variable

- A variable which contains the outcomes of a chance experiment
- "Quantifying the outcomes"
- Example X= (1 = Head, 0 = Tails)
- A variable that can take on different values in the population according to some "random" mechanism
- Discrete
  - Distinct values, countable
  - Year
- Continuous
  - Mass

Probability Distribution

- The probability distribution function or probability density function (PDF) of a random variable X means the values taken by that random variable and their associated probabilities.

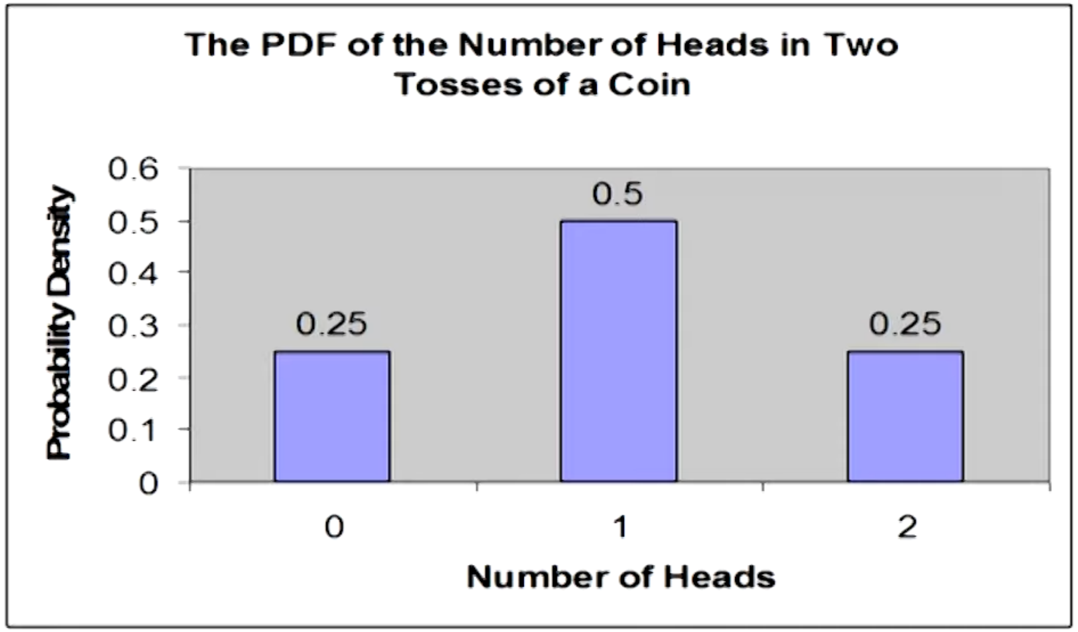
- PDF of a discrete r.v. (also known as PMF):

  Example 1: Let the r.v. X be the number of heads obtained in two tosses of a coin.

  Sample Space: {HH, HT, TH, TT}

PDF of Discrete r.v

Number of Heads (X):	0	1	2	sum
PDF ($P(X)$):	¼	½	¼	1



Probability Distribution for R.V – X

A probability distribution for a discrete random variable X:

x	–8	–3	–1	0	1	4	6
$P(X=x)$	0.13	0.15	0.17	0.20	0.15	0.11	0.09

Find

a. $P(X \le 0)$ $\boxed{0.65}$

b. $P(-3 \le X \le 1)$ $\boxed{0.67}$

Discrete Distribution –Example

Distribution of Daily Crises	
Number of Crises	Probability
0	0.37
1	0.31
2	0.18
3	0.09
4	0.04
5	0.01

Requirements for Discrete Probability Function

- Probabilities are between 0 and 1, inclusively

- Total of all probabilities equals 1

$$\boxed{0 \le P(X) \le 1 \quad \text{for all X}}$$

$$\boxed{\sum_{\text{over all x}} P(X) = 1}$$

Cumulative Distribution Function

- The CDF of a random variable X (defined as $F(X)$) is a graph associating all possible values, or the range of possible values with $P(X \leq x)$.
- CDFs always lie between 0 and 1 i.e., $0 \leq F(X_i) \leq 1$, Where $F(X_i)$ is the CDF.

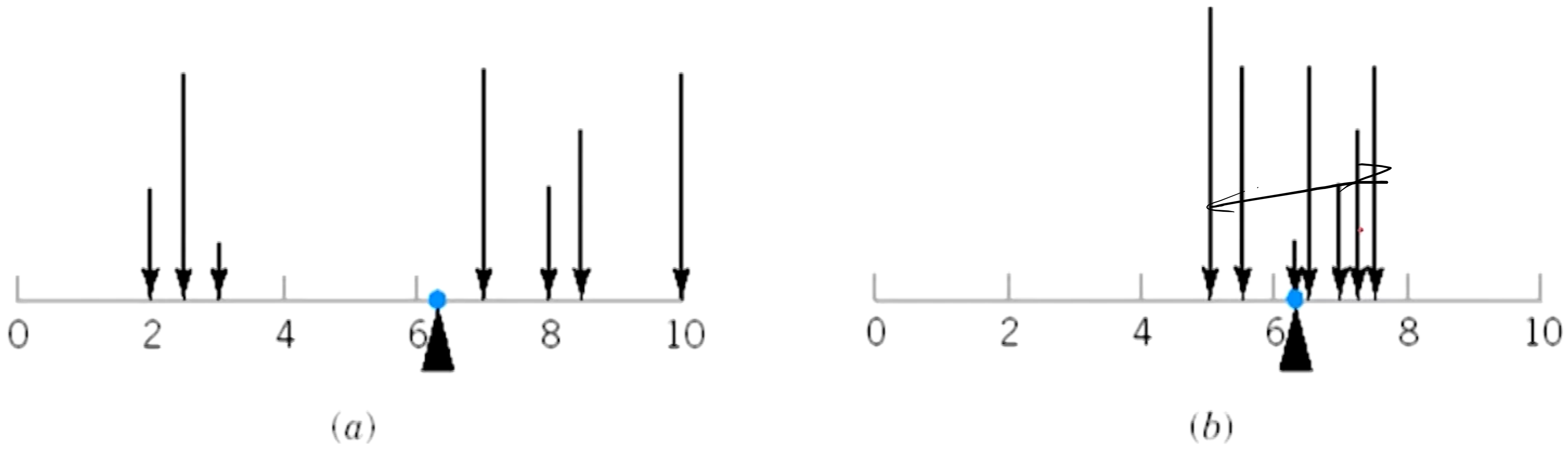
The Expected Value of X

Let X be a discrete rv with set of possible values D and pmf $p(x)$. The *expected value* or *mean value* of X, denoted

$$E(X) \text{ or } \mu_X, \text{ is}$$

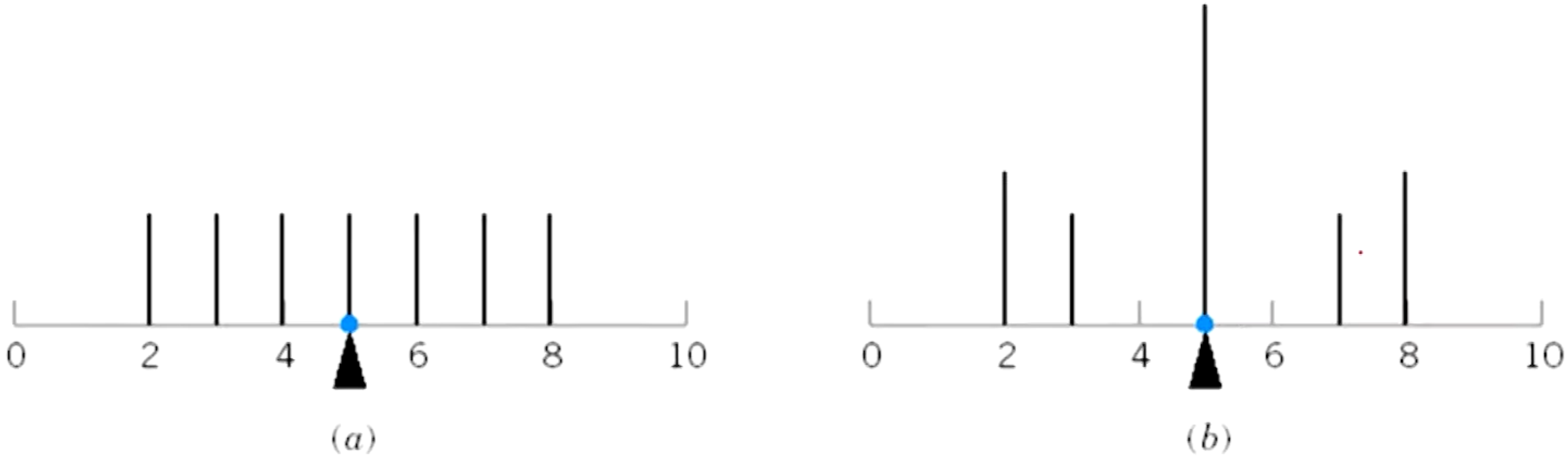
$$E(X) = \mu_X = \sum_{x \in D} x \cdot p(x)$$

Mean and Variance of a Discrete R.V



A probability distribution can be viewed as a loading with the mean equal to the balance point. Parts (a) and (b) illustrate equal means, but Part (a) illustrates a larger variance.

Mean and Variance of a Discrete R.V



The probability distribution illustrated in Parts (a) and (b) differ even though they have equal means and equal variances.

Example-Expected Value

- Use the data below to find out the expected number of credit cards that a customer to a retail outlet will possess.

$x = \#$ credit cards

x	$P(x = X)$
0	0.08
1	0.28
2	0.38
3	0.16
4	0.06
5	0.03
6	0.01

$$E(X) = x_1 p_1 + x_2 p_2 + \ldots + x_n p_n$$

$$= 0(.08) + 1(.28) + 2(.38) + 3(.16) + 4(.06) + 5(.03) + 6(.01)$$

$$= 1.97$$

About 2 credit cards

The Variance and Standards Deviation

Let X have pmf p(x), and expected value μ

Then the variance of X, denoted V(X)

(or σ_X^2 or σ^2), is

$$V(X) = \sum_D (x-\mu)^2 \cdot p(x) = E[(X-\mu)^2]$$

The standard deviation (SD) of X is

$$\sigma_X = \sqrt{\sigma_X^2}$$

Problem

The quiz scores for a particular student are given below:

22, 25, 20, 18, 12, 20, 24, 20, 20, 25, 24, 25, 18

Find the variance and standard deviation.

Value	12	18	20	22	24	25
Frequency	1	2	4	1	2	3
Probability	.08	.15	.31	.08	.15	.23

$\mu = 21$

$$V(X) = p_1(x_1 - \mu)^2 + p_2(x_2 - \mu)^2 + \ldots + p_n(x_n - \mu)^2$$

$$\sigma = \sqrt{V(X)}$$

$$V(X) = .08(12-21)^2 + .15(18-21)^2 + .31(20-21)^2$$
$$+ .08(22-21)^2 + .15(24-21)^2 + .23(25-21)^2$$

$$\boxed{V(X) = 13.25}$$

$$\sigma = \sqrt{V(X)} \quad \boxed{= \sqrt{13.25} \approx 3.64}$$

Shortcut Formula for Variance

$$V(X) = \sigma^2 = \left[\sum_D x^2 \cdot p(x)\right] - \mu^2$$

$$= E\left(X^2\right) - \left[E(X)\right]^2$$

Mean of a Discrete Distribution

$$\boxed{\mu = E(X) = \sum X \cdot P(X)}$$

X	P(X)	X.P(X)
-1	.1	-.1
0	.2	.0
1	.4	.4
2	.2	.4
3	.1	.3
		1.0

Variance and Standard Deviation of Discrete Distribution

$$\boxed{\sigma^2 = \sum (X-\mu)^2 \cdot P(X) = 1.2}$$

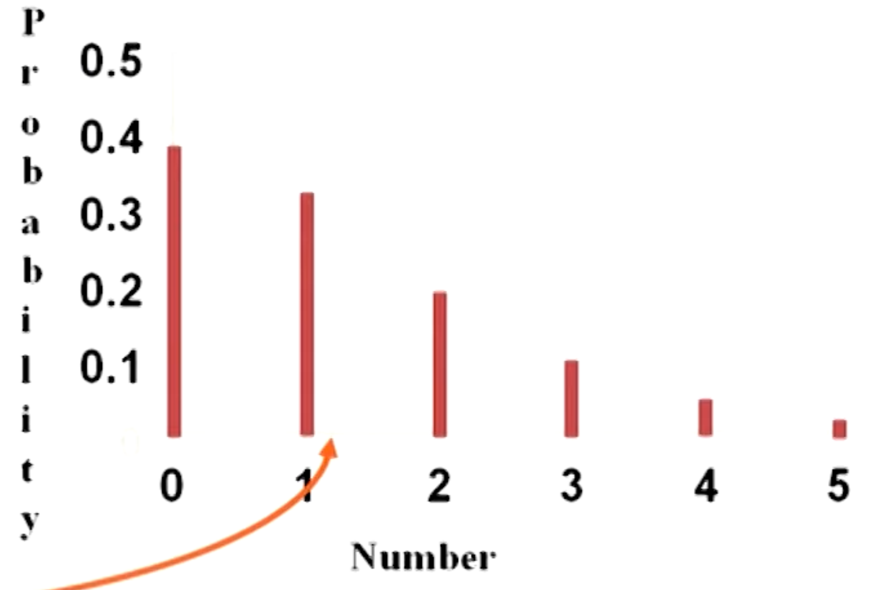
$$\boxed{\sigma = \sqrt{\sigma^2} = \sqrt{1.2} \cong 1.10}$$

X	P(X)	$X - \mu$	$(X-\mu)^2$	$(X-\mu)^2 \cdot P(X)$
-1	1	-2	4	.4
0	.2	-1	1	.2
1	.4	0	0	.0
2	.2	1	1	.2
3	.1	2	4	.4
				1.2

Mean of the Data Example

$$\mu = E(X) = \sum X \cdot P(X) = 1.15$$

X	P(X)	X•P(X)
0	.37	.00
1	.31	.31
2	.18	.36
3	.09	.27
4	.04	.16
5	.01	.05
		1.15



Properties of Expected Value

1. $E(b) = b, b$ is a constant.

2. $E(X + Y) = E(X) + E(Y)$.

3. $E\left(X/Y\right) \neq \dfrac{E(X)}{E(Y)}$.

4. $E(XY) \neq E(X)E(Y)$ unless they are indpendendent.

5. $E(aX) = aE(X)$, a constant.

6. $E(aX + b) = aE(X) + b$, a and b are constants.

Properties of Variance

1. $Var(constant) = 0$
2. If X and Y are two independent random variables, then

   $Var(X + Y) = Var(X) + Var(Y)$ and

   $Var(X - Y) = Var(X) + Var(Y)$
3. If b is a constant then $Var(b+X) = Var(X)$
4. If a is a constant then $Var(aX) = a^2Var(X)$
5. If a and b are constants then $Var(aX+b) = a^2Var(X)$
6. If X and Y are two independent random variables and a and b are constants then $Var(aX+bY) = a^2Var(X) + b^2Var(Y)$

Covariance

Covariance: For two discrete random variables X and Y with $E(X) = \mu_x$ and $E(Y) = \mu_y$, the covariance between X and Y is defined as Cov(XY) $= \sigma_{xy} = E(X - \mu_x)\, E(Y - \mu_y) = E(XY) - \mu_x \mu_y$.

Covariance

- In general, the covariance between two random variables can be positive or negative.
- If two random variables move in the same direction, then the covariance will be positive, if they move in the opposite direction the covariance will be negative.

Properties:

1. If X and Y are independent random variables, their covariance is zero. Since $E(XY) = E(X)E(Y)$
2. Cov(XX) = Var(X)
3. Cov(YY) = Var(Y)

Covariance

Covariance between two random variables:

- $cov(X,Y) > 0$ X and Y are positively correlated
- $cov(X,Y) < 0$ X and Y are inversely correlated
- $cov(X,Y) = 0$ X and Y are independent

Correlation Coefficient

- The covariance tells the sign but not the magnitude about how strongly the variables are positively or negatively related. The correlation coefficient provides such measure of how strongly the variables are related to each other.
- For two random variables X and Y with $E(X) = \mu_x$ and $E(Y) = \mu_y$, the correlation coefficient is defined as

$$\rho_{xy} = \frac{Cov(XY)}{\sigma_x \sigma_y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y}$$

Continuous Random Variables

Probability Density Function

- The probability distribution or simply distribution of a random variable X is a description of the set of the probabilities associated with the possible values for X.

The **probability density function** (or pdf) $f(x)$ of a continuous random variable is used to determine probabilities from areas as follows:

$$P(a < X < b) = \int_a^b f(x)\,dx \qquad (3\text{-}2)$$

The properties of the pdf are

(1) $f(x) \geq 0$

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

Probability Density Function

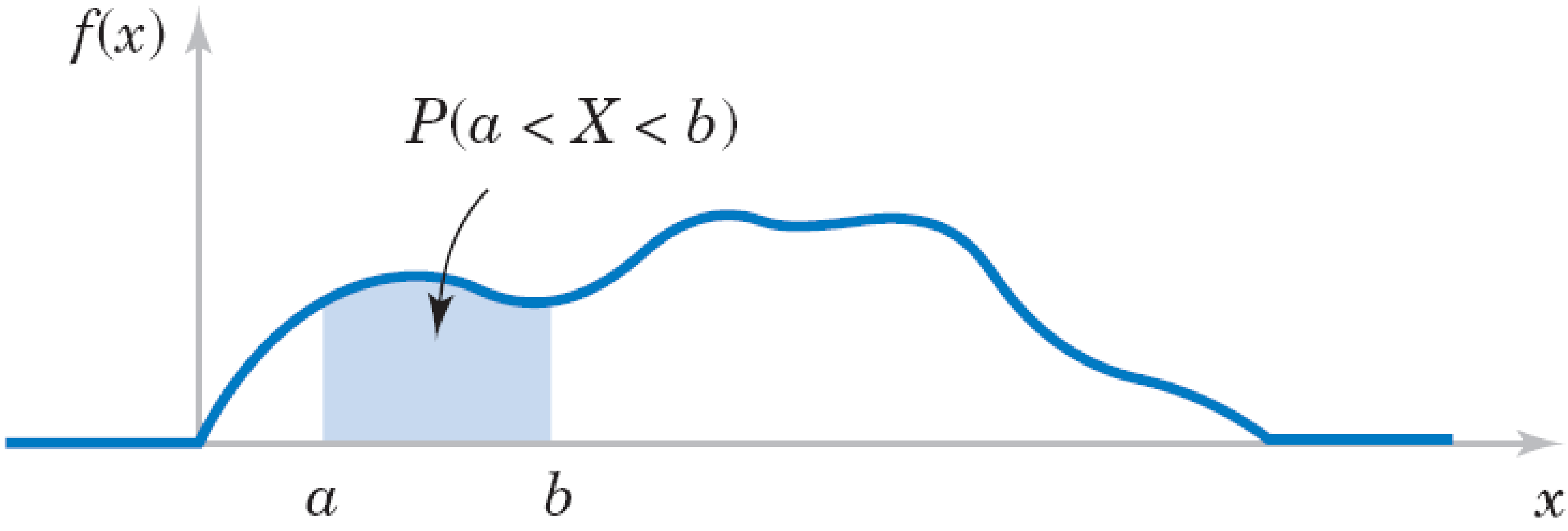


Figure 3-6 Probability determined from the area under $f(x)$.

Probability Density Function

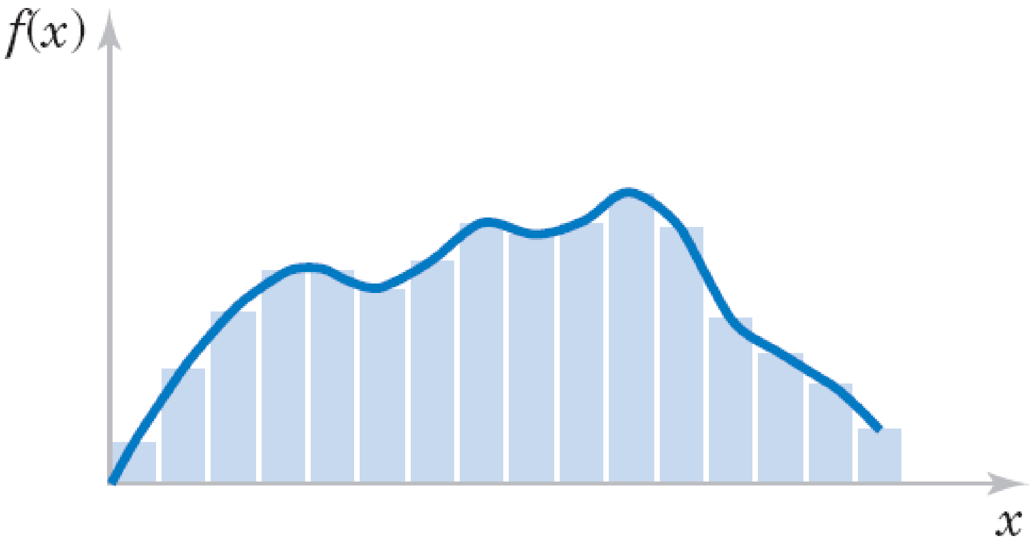


Figure 3-7 A histogram approximates a probability density function. The area of each bar equals the relative frequency of the interval. The area under $f(x)$ over any interval equals the probability of the interval.

Probability Density Function

If X is a continuous random variable, for any x_1 and x_2,

$$P(x_1 \leq X \leq x_2) = P(x_1 < X \leq x_2) = P(x_1 \leq X < x_2) = P(x_1 < X < x_2)$$

Probability Density Function

Let the continuous random variable X denote the distance in micrometers from the start of a track on a magnetic disk until the first flaw. Historical data show that the distribution of X can be modeled by a pdf $f(x) = \frac{1}{2000} e^{-x/2000}$, $x \geq 0$. For what proportion of disks is the distance to the first flaw greater than 1000 micrometers?

Solution. The density function and the requested probability are shown in Fig. 3-9.
Now,

$$P(X > 1000) = \int_{1000}^{\infty} f(x)\, dx = \int_{1000}^{\infty} \frac{e^{-x/2000}}{2000}\, dx = -e^{-x/2000} \Big|_{1000}^{\infty} = e^{-1/2} = 0.607$$

What proportion of parts is between 1000 and 2000 micrometers?

Solution. Now,

$$P(1000 < X < 2000) = \int_{1000}^{2000} f(x)\, dx = -e^{-x/2000} \Big|_{1000}^{2000} = e^{-1/2} - e^{-1} = 0.239$$

Because the total area under $f(x)$ equals 1, we can also calculate $P(X < 1000) = 1 - P(X > 1000) = 1 - 0.607 = 0.393$.

Cumulative Distribution Function

The **cumulative distribution function** (or cdf) of a continuous random variable X with probability density function $f(x)$ is

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(u)\,du$$

for $-\infty < x < \infty$.

Mean and Variance

Suppose X is a continuous random variable with pdf $f(x)$. The **mean** or **expected value** of X, denoted as μ or $E(X)$, is

$$\mu = E(X) = \int_{-\infty}^{\infty} x f(x)\,dx \tag{3-3}$$

The **variance** of X, denoted as $V(X)$ or σ^2, is

$$\sigma^2 = V(X) = \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,dx = \int_{-\infty}^{\infty} x^2 f(x)\,dx - \mu^2$$

The **standard deviation** of X is σ.

Discrete vs Continuous Variable

Discrete variable

- Countable support
- Probability mass function
- Probabilities assigned to single values
- Each possible value has strictly positive probability

Continuous variable

- Uncountable support
- Probability density function
- Probabilities assigned to intervals of values
- Each possible value has zero probability

Some Special Distribution

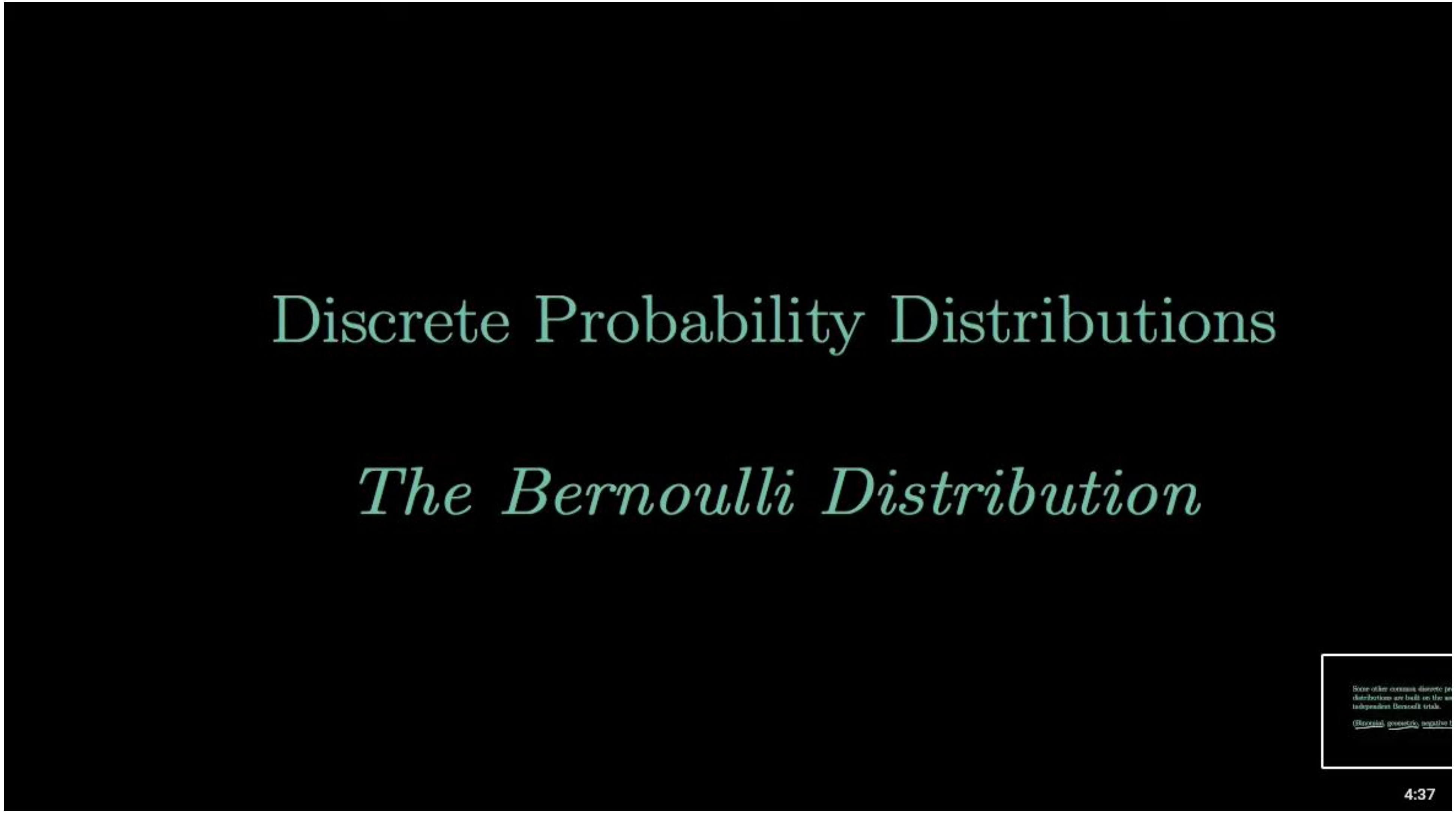
Discrete

- Bernoulli
- Geometric
- Binomial
- Poisson

Continuous

- Uniform
- Normal
- Exponential

The Bernoulli Distribution



The Bernoulli Distribution

Suppose we have

- A single trial.
- The trial can result in one of two possible outcomes, labelled success and failure.
- $P(\text{Success}) = p$
- $P(\text{Failure}) = 1 - p$

The Bernoulli Distribution

Let $X = 1$ if a success occurs, and
$X = 0$ if a failure occurs.

Then X has a Bernoulli distribution:

$$P(X = x) = p^x (1 - p)^{1-x}$$

The Bernoulli Distribution

Approximately 1 in 200 American adults are lawyers.

One American adult is randomly selected.

What is the distribution of the number of lawyers?

The Bernoulli Distribution

Approximately 1 in 200 American adults are lawyers.

One American adult is randomly selected.

What is the distribution of the number of lawyers? Bernoulli with $p = \frac{1}{200}$

X

$$P(X = x) = \left(\frac{1}{200}\right)^{x} \left(1 - \frac{1}{200}\right)^{1-x}$$

for $x = 0, 1$

$$P(X = 1) = \frac{1}{200} \qquad P(X = 0) = \frac{199}{200}$$

The Bernoulli Distribution

Some other common discrete probability distributions are built on the assumption of independent Bernoulli trials.

(Binomial, geometric, negative binomial)

Geometric Distribution

The geometric distribution is the distribution of the number of trials needed to get the first success in repeated Bernoulli trials.

Geometric Distribution

Suppose:

- There are independent trials.
- Each trial results in one of two possible outcomes, labelled success and failure.

Geometric Distribution

- $P(\text{Success}) = p$, and this stays constant from trial to trial.
- $P(\text{Failure}) = 1 - p$.
- X represents the number of trials needed to get the first success.

Geometric Distribution

For the first success to occur on the xth trial:

$(1-p)^{x-1}$

1. The first $x - 1$ trials must be failures.
2. The xth trial must be a success. p

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

$$P(X = x) = (1 - p)^{x-1}p$$

for $x = 1, 2, 3, \ldots$

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

Geometric Distribution

In a large population of adults, 30% have received CPR training.

$p = 0.3$

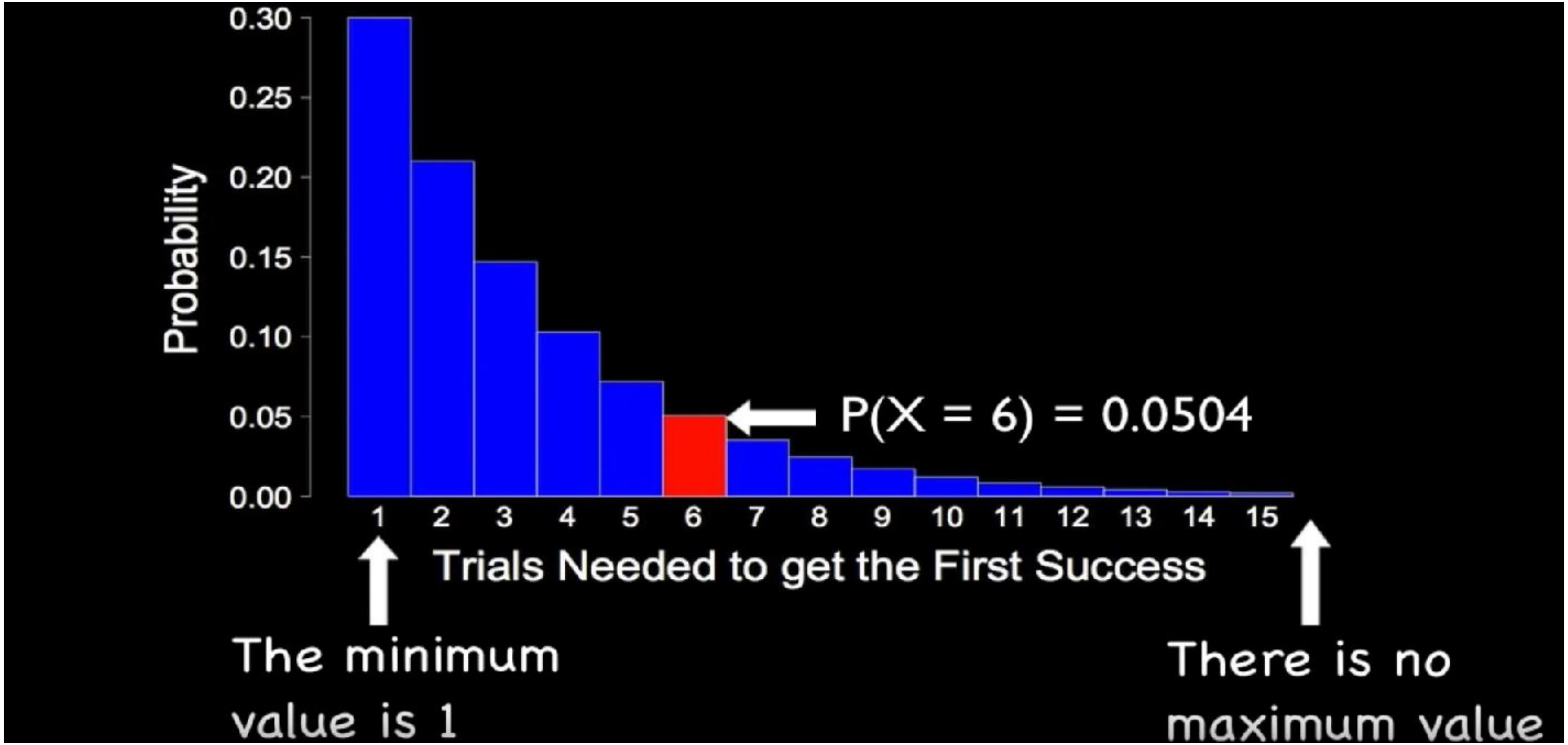
If adults from this population are randomly selected, what is the probability that the 6th person sampled is the first that has received CPR training?

$$P(X = x) = (1 - p)^{x-1} p$$

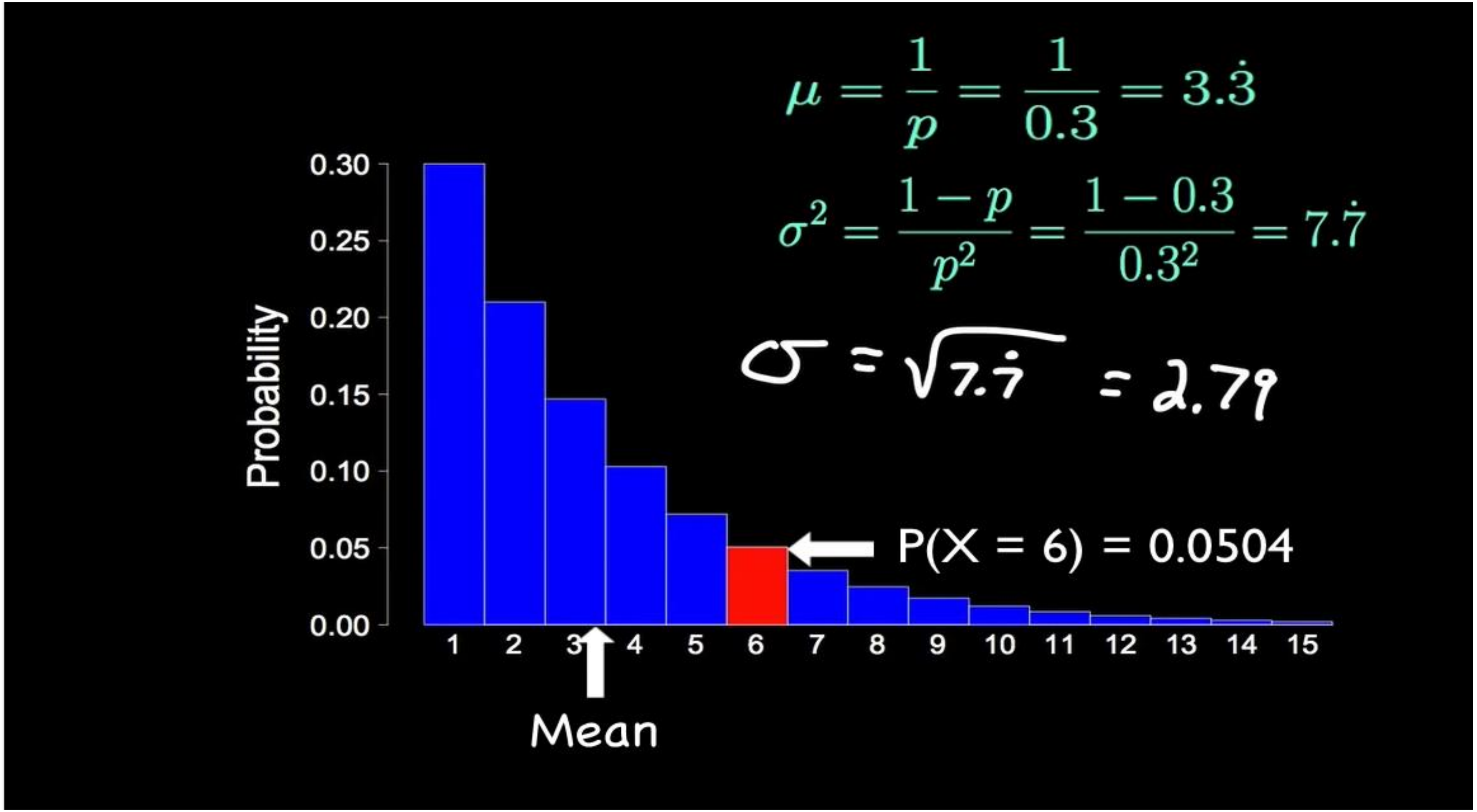
$$P(X=6) = (1-0.3)^5 0.3$$

$$= 0.0504$$

Geometric Distribution



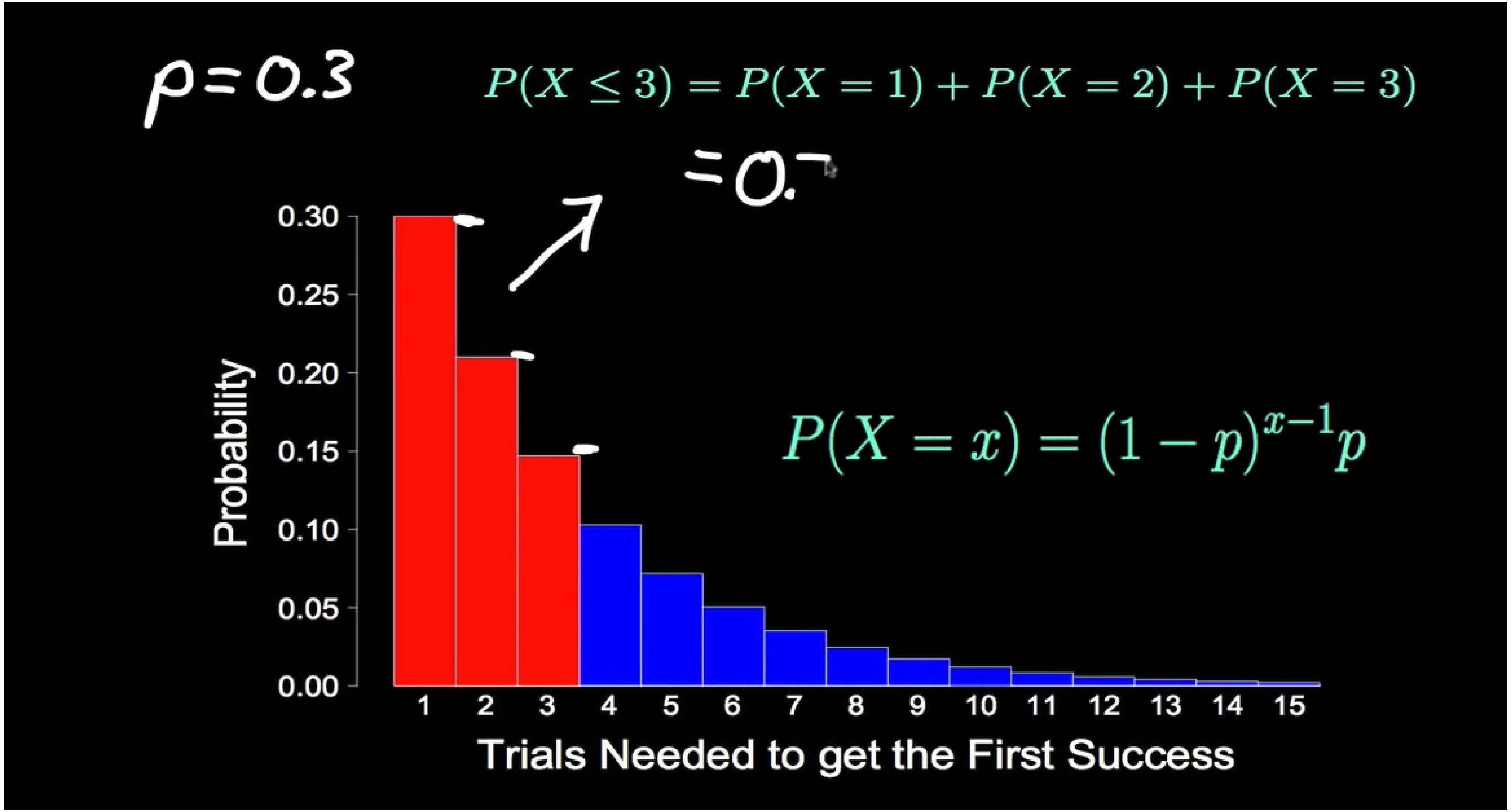
Geometric Distribution



Geometric Distribution

What is the probability that the first person trained in CPR occurs on or before the 3rd person sampled?

Geometric Distribution



Geometric Distribution

The cumulative distribution function for the geometric distribution:

$$F(x) = P(X \leq x) = 1 - (1-p)^x$$

for $x = 1, 2, 3, \ldots$

Binomial Distribution

Binomial Distribution -Assumptions

- Experiment involves n identical trials
- Each trial has exactly two possible outcomes: success and failure
- Each trial is independent of the previous trials
- p is the probability of a success on any one trial

  q = (1-p) is the probability of a failure on any one trial
- p and q are constant throughout the experiment
- X is the number of successes in the n trials

Binomial Distribution

- Probability function

$$P(X)=\frac{n!}{X!(n-X)!}p^{X}\cdot q^{n-X} \quad \text{for} \quad 0\leq X\leq n$$

- Mean value

$$\mu=n\cdot p$$

- Variance and standard deviation

$$\sigma^{2}=n\cdot p\cdot q$$

$$\sigma=\sqrt{\sigma^{2}}=\sqrt{n\cdot p\cdot q}$$

Binomial Table

SELECTED VALUES FROM THE BINOMIAL PROBABILITY TABLE

EXAMPLE: $n = 10$, $x = 3$, $p = .40$; $f(3) = .2150$

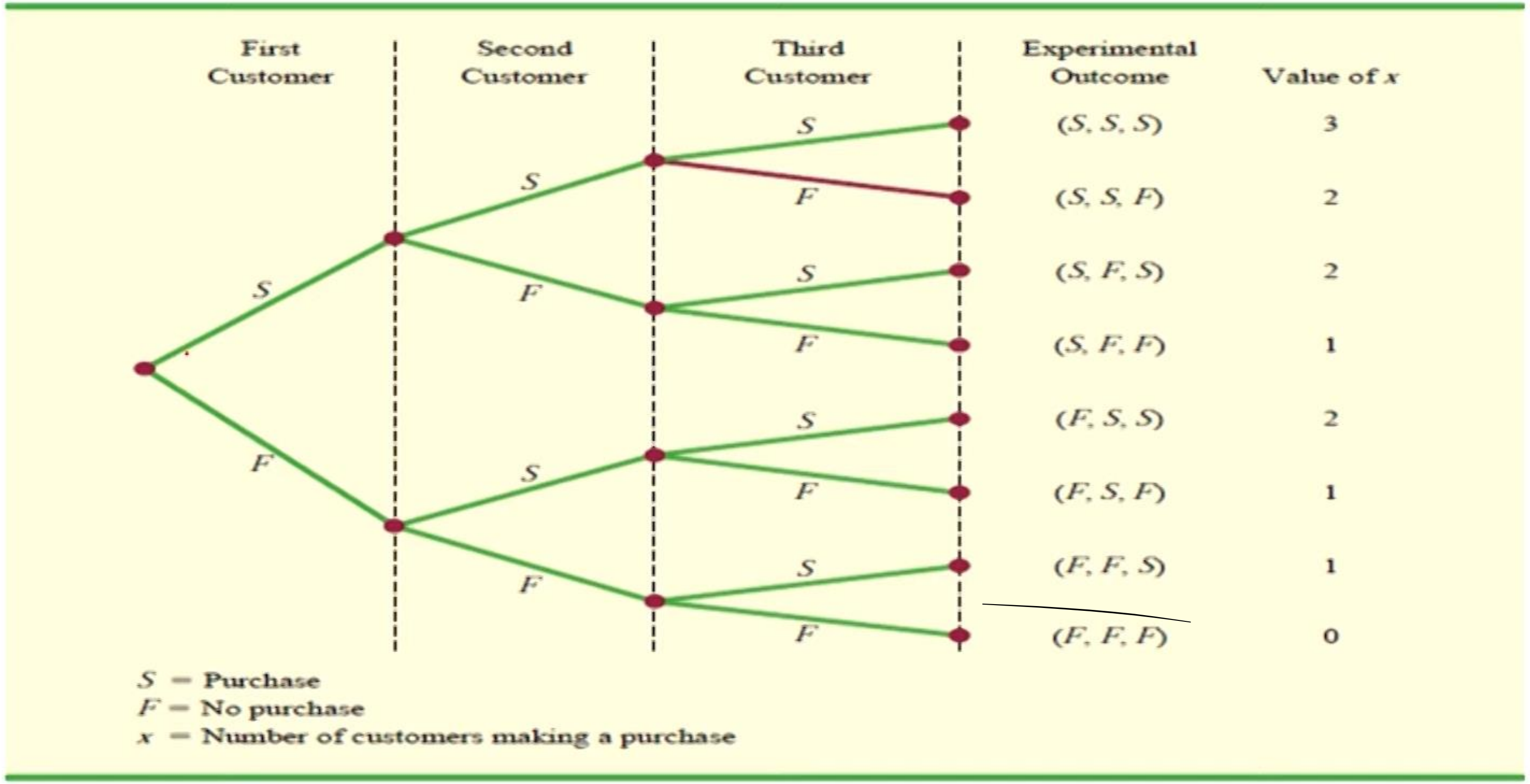
		p									
n	x	.05	.10	.15	.20	.25	.30	.35	.40	.45	.50
9	0	.6302	.3874	.2316	.1342	.0751	.0404	.0207	.0101	.0046	.0020
	1	.2985	.3874	.3679	.3020	.2253	.1556	.1004	.0605	.0339	.0176
	2	.0629	.1722	.2597	.3020	.3003	.2668	.2162	.1612	.1110	.0703
	3	.0077	.0446	.1069	.1762	.2336	.2668	.2716	.2508	.2119	.1641
	4	.0006	.0074	.0283	.0661	.1168	.1715	.2194	.2508	.2600	.2461
	5	.0000	.0008	.0050	.0165	.0389	.0735	.1181	.1672	.2128	.2461
	6	.0000	.0001	.0006	.0028	.0087	.0210	.0424	.0743	.1160	.1641
	7	.0000	.0000	.0000	.0003	.0012	.0039	.0098	.0212	.0407	.0703
	8	.0000	.0000	.0000	.0000	.0001	.0004	.0013	.0035	.0083	.0176
	9	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0008	.0020
10	0	.5987	.3487	.1969	.1074	.0563	.0282	.0135	.0060	.0025	.0010
	1	.3151	.3874	.3474	.2684	.1877	.1211	.0725	.0403	.0207	.0098
	2	.0746	.1937	.2759	.3020	.2816	.2335	.1757	.1209	.0763	.0439
	3	.0105	.0574	.1298	.2013	.2503	.2668	.2522	**.2150**	.1665	.1172
	4	.0010	.0112	.0401	.0881	.1460	.2001	.2377	.2508	.2384	.2051
	5	.0001	.0015	.0085	.0264	.0584	.1029	.1536	.2007	.2340	.2461
	6	.0000	.0001	.0012	.0055	.0162	.0368	.0689	.1115	.1596	.2051
	7	.0000	.0000	.0001	.0008	.0031	.0090	.0212	.0425	.0746	.1172
	8	.0000	.0000	.0000	.0001	.0004	.0014	.0043	.0106	.0229	.0439
	9	.0000	.0000	.0000	.0000	.0000	.0001	.0005	.0016	.0042	.0098
	10	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0003	.0010

Binomial Distribution

- Let us consider the purchase decisions of the next three customers who enter a store.
- On the basis of past experience, the store manager estimates the probability that any one customer will make a purchase is .30.
- What is the probability that two of the next three customers will make a purchase?

Binomial Distribution

Tree diagram for the Martin clothing store problem



Binomial Distribution

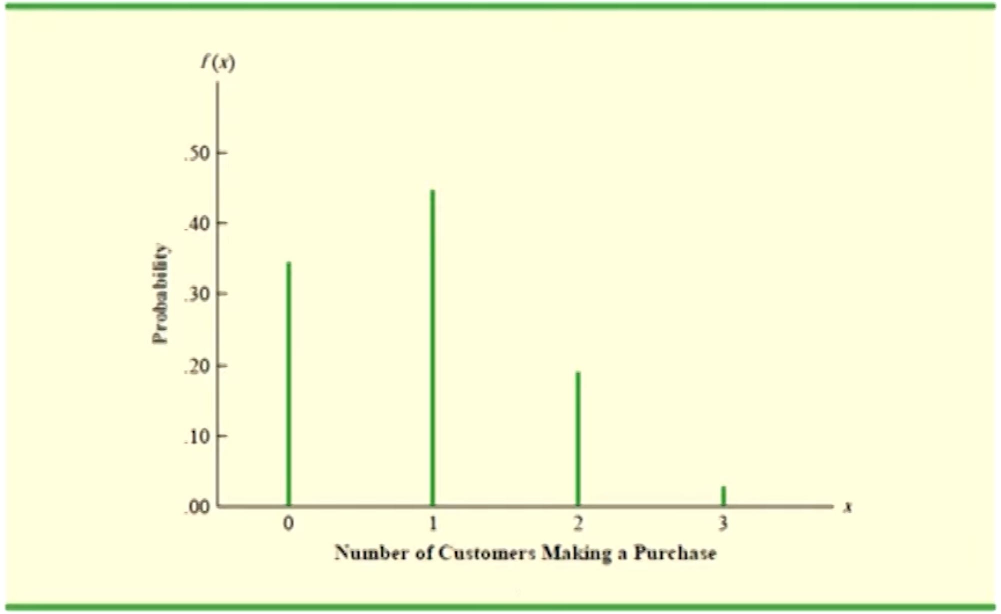
Trial Outcomes

Trial Outcomes: 1st Customer	2nd Customer	3rd Customer	Experimental Outcome	Probability of Experimental Outcome
Purchase	Purchase	No purchase	(S, S, F)	$pp(1-p) = p^2(1-p)$ $= (.30)^2(.70) = .063$
Purchase	No purchase	Purchase	(S, F, S)	$p(1-p)p = p^2(1-p)$ $= (.30)^2(.70) = .063$
No purchase	Purchase	Purchase	(F, S, S)	$(1-p)pp = p^2(1-p)$ $= (.30)^2(.70) = .063$

Binomial Distribution

Graphical representation of the probability distribution for the number of customers making a purchase

x	P(x)
0	0.7 x 0.7 x 0.7=0.343
1	0.3x0.7x07+ 0.7x0.3x0.7+ 0.7x0.7x0.3 = 0.441
2	0.189
3	0.027



Binomial Distribution

Mean and Variance

- Suppose that for the next month the Clothing Store forecasts 1000 customers will enter the store.
- What is the expected number of customers who will make a purchase?
- The answer is $\mu = np =$ (1000)(.3) = 300.
- For the next 1000 customers entering the store, the variance and standard deviation for the number of customers who will make a purchase are

$$\sigma^2 = np(1-p) = 1000(.3)(.7) = 210$$
$$\sigma = \sqrt{210} = 14.49$$

Poisson Distribution

- Describes discrete occurrences over a continuum or interval
- A discrete distribution
- Describes rare events
- Each occurrence is independent any other occurrences.
- The number of occurrences in each interval can vary from zero to infinity.
- The expected number of occurrences must hold constant throughout the experiment.

Poisson Distribution: Application

- **Arrivals at queuing systems**
  - airports -- people, airplanes, automobiles, baggage
  - banks -- people, automobiles, loan applications
  - computer file servers -- read and write operations

- **Defects in manufactured goods**
  - number of defects per 1,000 feet of extruded copper wire
  - number of blemishes per square foot of painted surface
  - number of errors per typed page

Poisson Distribution

- Probability function

$$P(X) = \frac{\lambda^{X} e^{-\lambda}}{X!} \quad \text{for} \quad X = 0, 1, 2, 3, \ldots$$

where:

$\lambda = long-run\ average$

$e = 2.718282\ldots$ (the base of natural logarithms)

Mean value	Variance	Standard deviation
$\boxed{\lambda}$	$\boxed{\lambda}$	$\boxed{\sqrt{\lambda}}$

Poisson Distribution- Example

$\lambda = 3.2$ customers/4 minutes

X = 10 customers/8 minutes

Adjusted λ

λ=6.4 customers/8 minutes

$$P(X) = \frac{\lambda^{X} e^{-\lambda}}{X!}$$

$$P(X=10) = \frac{6.4^{10} e^{-6.4}}{10!} = 0.0528$$

$\lambda = 3.2$ customers/4 minutes

X = 6 customers/8 minutes

Adjusted λ

λ=6.4 customers/8 minutes

$$P(X) = \frac{\lambda^{X} e^{-\lambda}}{X!}$$

$$P(X=6) = \frac{6.4^{6} e^{-6.4}}{6!} = 0.1586$$

Poisson Distribution- Example

Poisson Probability Table

Example: μ = 10, x = 5; f (5) = .0378

x	μ = 9.1	9.2	9.3	9.4	9.5	9.6	9.7	9.8	9.9	10
0	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0000
1	.0010	.0009	.0009	.0008	.0007	.0007	.0006	.0005	.0005	.0005
2	.0046	.0043	.0040	.0037	.0034	.0031	.0029	.0027	.0025	.0023
3	.0140	.0131	.0123	.0115	.0107	.0100	.0093	.0087	.0081	.0076
4	.0319	.0302	.0285	.0269	.0254	.0240	.0226	.0213	.0201	.0189
5	.0581	.0555	.0530	.0506	.0483	.0460	.0439	.0418	.0398	.0378
6	.0881	.0851	.0822	.0793	.0764	.0736	.0709	.0682	.0656	.0631
7	.1145	.1118	.1091	.1064	.1037	.1010	.0982	.0955	.0928	.0901
8	.1302	.1286	.1269	.1251	.1232	.1212	.1191	.1170	.1148	.1126
9	.1317	.1315	.1311	.1306	.1300	.1293	.1284	.1274	.1263	.1251
10	.1198	.1210	.1219	.1228	.1235	.1241	.1245	.1249	.1250	.1251
11	.0991	.1012	.1031	.1049	.1067	.1083	.1098	.1112	.1125	.1137
12	.0752	.0776	.0799	.0822	.0844	.0866	.0888	.0908	.0928	.0948
13	.0526	.0549	.0572	.0594	.0617	.0640	.0662	.0685	.0707	.0729
14	.0342	.0361	.0380	.0399	.0419	.0439	.0459	.0479	.0500	.0521
15	.0208	.0221	.0235	.0250	.0265	.0281	.0297	.0313	.0330	.0347
16	.0118	.0127	.0137	.0147	.0157	.0168	.0180	.0192	.0204	.0217
17	.0063	.0069	.0075	.0081	.0088	.0095	.0103	.0111	.0119	.0128
18	.0032	.0035	.0039	.0042	.0046	.0051	.0055	.0060	.0065	.0071
19	.0015	.0017	.0019	.0021	.0023	.0026	.0028	.0031	.0034	.0037
20	.0007	.0008	.0009	.0010	.0011	.0012	.0014	.0015	.0017	.0019
21	.0003	.0003	.0004	.0004	.0005	.0006	.0006	.0007	.0008	.0009
22	.0001	.0001	.0002	.0002	.0002	.0002	.0003	.0003	.0004	.0004
23	.0000	.0001	.0001	.0001	.0001	.0001	.0001	.0001	.0002	.0002
24	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0001	.0001

Poisson Distribution- Example

One nanogram of Plutonium-239 will have an average of 2.3 radioactive decays per second, and the number of decays will follow a Poisson distribution.

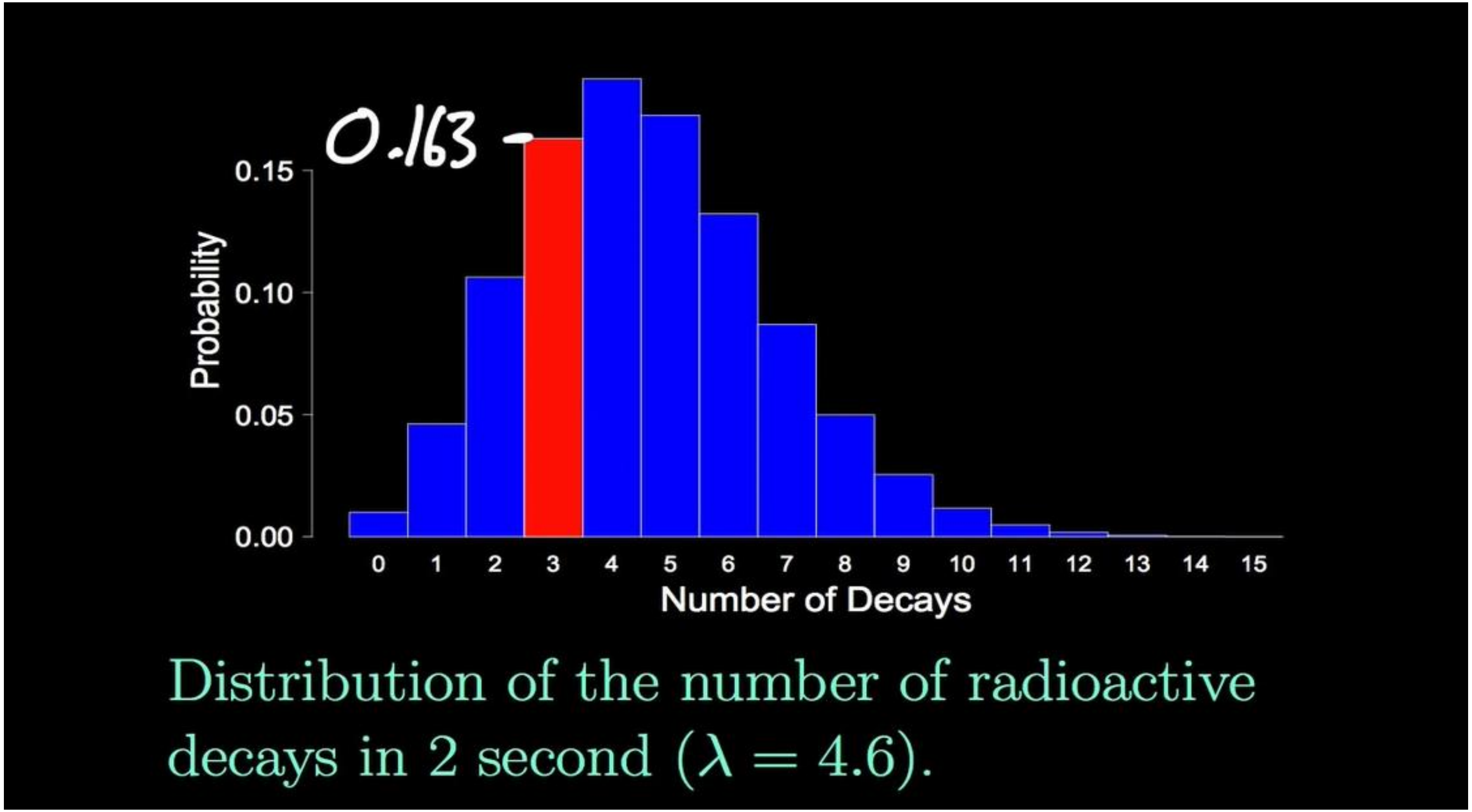
What is the probability that in a 2 second period there are exactly 3 radioactive decays?

Let X represent the number of decays in a 2 second period.

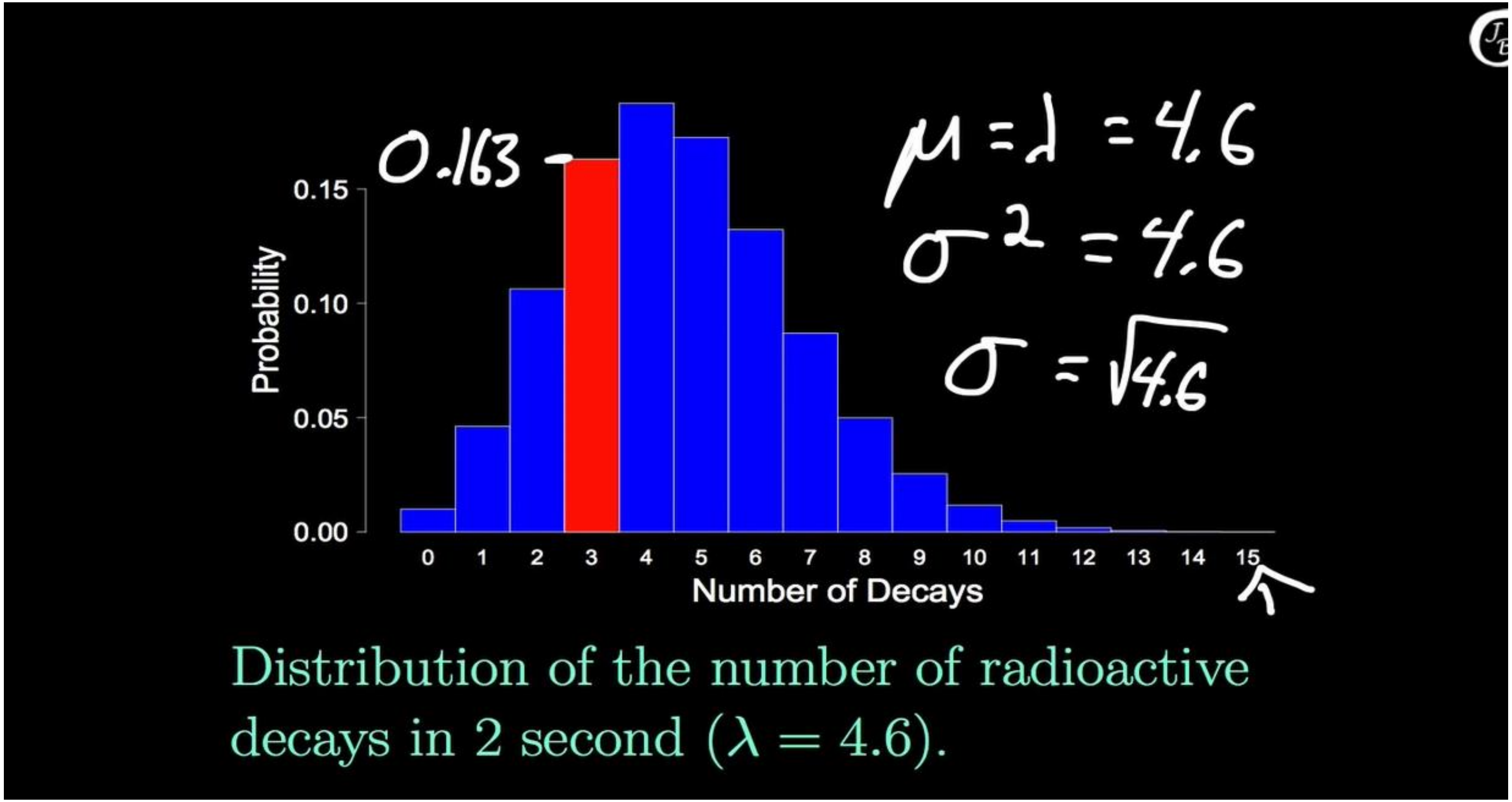
$$\lambda = 2.3 \times 2 = 4.6$$

$$P(X=3) = \frac{\lambda^x e^{-\lambda}}{x!} = \frac{4.6^3 e^{-4.6}}{3!} = 0.163$$

Poisson Distribution- Example



Poisson Distribution- Example



Poisson Distribution- Example

$$P(X \leq 3) = P(X = 0) + P(X = 1) + P(X = 2) + P(X = 3)$$

$$= \frac{4.6^0 e^{-4.6}}{0!} + \frac{4.6^1 e^{-4.6}}{1!} + \frac{4.6^2 e^{-4.6}}{2!} + \frac{4.6^3 e^{-4.6}}{3!}$$

$$= 0.010 + 0.046 + 0.106 + 0.163$$

$$= 0.326$$

Continuous Distribution

Continuous Probability Distribution

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

Continuous Probability Distribution

Continuous

- Uniform
- Normal
- Exponential

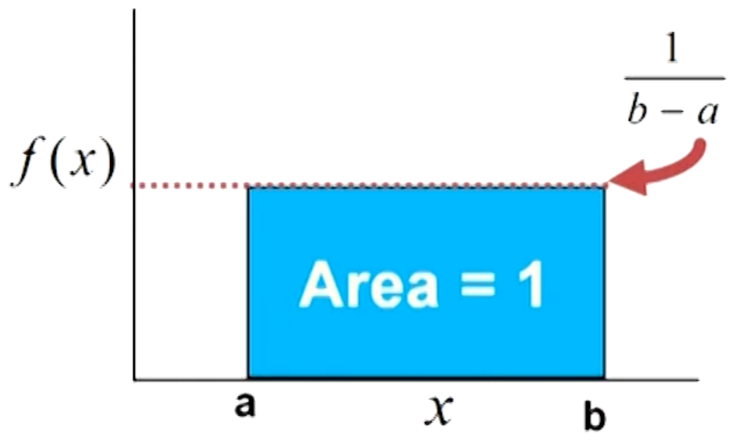
The Uniform Distribution

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

- Because of its shape it is also called a rectangular distribution

The Uniform Distribution

$$f(x) = \begin{cases} \dfrac{1}{b-a} & \text{for} \quad a \le x \le b \\ 0 & \text{for} \quad \text{all other valu es} \end{cases}$$



The Uniform Distribution: Mean and Std

Mean

$$\mu = \frac{a + b}{2}$$

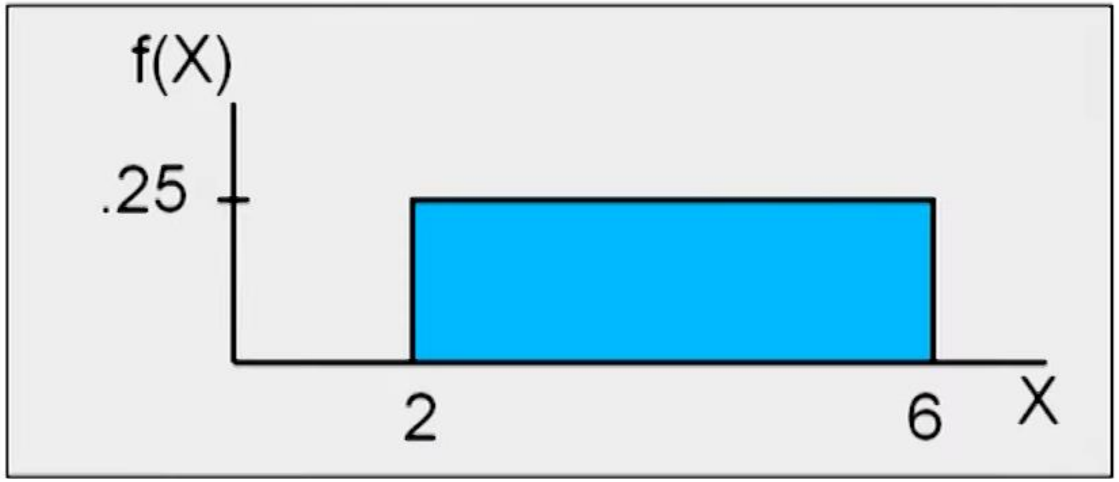
Standard Deviation

$$\sigma = \frac{b - a}{\sqrt{12}}$$

The Uniform Distribution: Mean and Std

Example: Uniform probability distribution over the range $2 \leq X \leq 6$:

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

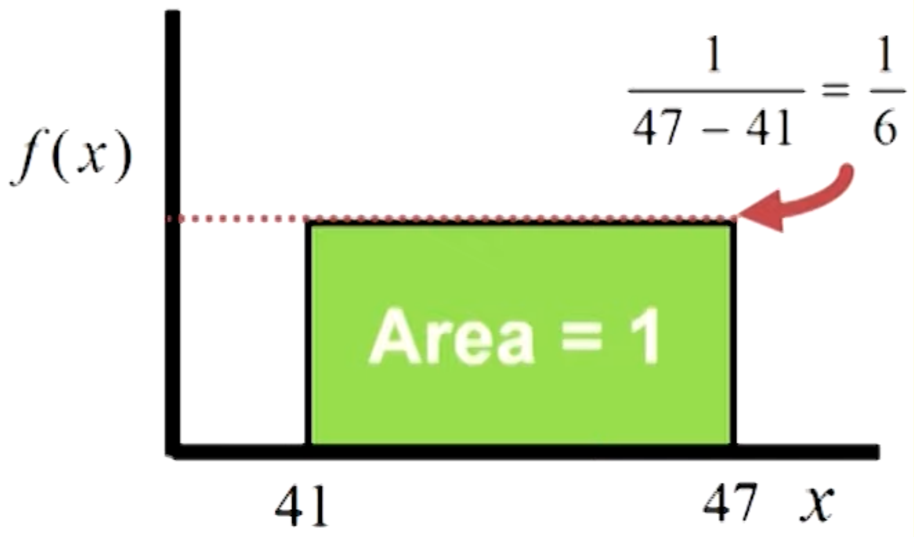


$$\mu = \frac{a+b}{2} = \frac{2+6}{2} = 4$$

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

The Uniform Distribution –Example

$$f(x) = \begin{cases} \dfrac{1}{47-41} & for \quad 41 \le x \le 47 \\[2ex] 0 & for \quad \text{all other values} \end{cases}$$



The Uniform Distribution – Mean and Std

Mean

$$\mu = \frac{a + b}{2}$$

Mean

$$\mu = \frac{41+47}{2} = \frac{88}{2} = 44$$

Standard Deviation

$$\sigma = \frac{b-a}{\sqrt{12}}$$

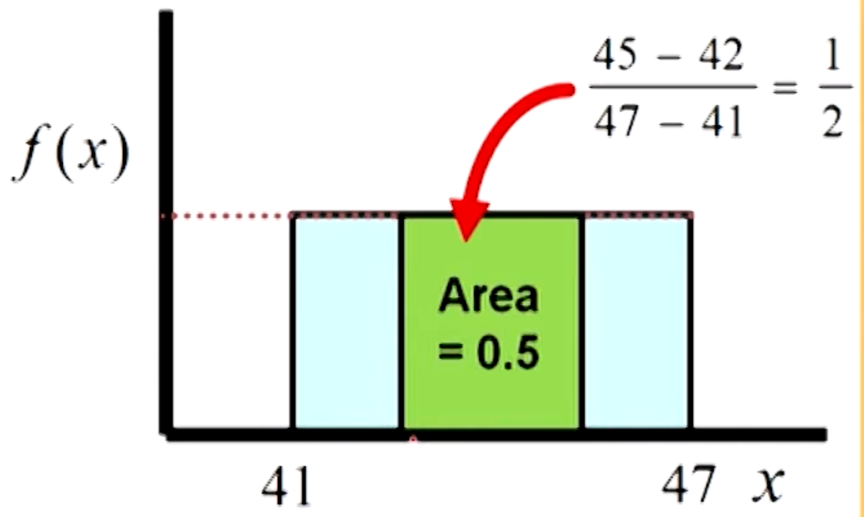
Standard Deviation

$$\sigma = \frac{47-41}{\sqrt{12}} = \frac{6}{3.464} = 1.732$$

The Uniform Distribution Probability

$$P(x_1 \le X \le x_2) = \frac{x_2 - x_1}{b - a}$$

$$P(42 \le X \le 45) = \frac{45-42}{47-41} = \frac{1}{2}$$



The Uniform Distribution- Example

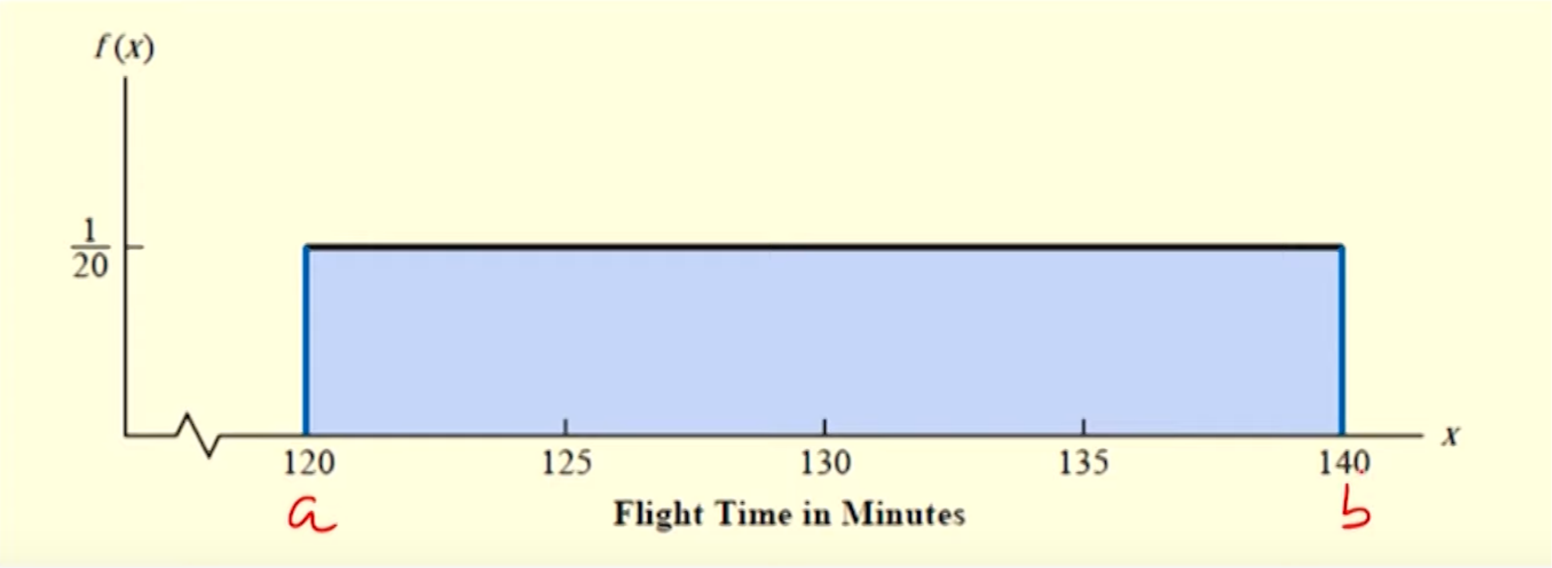
- Consider the random variable *x* representing the flight time of an airplane traveling from Delhi to Mumbai.
- Suppose the flight time can be any value in the interval from 120 minutes to 140 minutes.
- Because the random variable *x* can assume any value in that interval, *x* is a continuous rather than a discrete random variable

The Uniform Distribution- Example

- Let us assume that sufficient actual flight data are available to conclude that the probability of a flight time within any 1-minute interval is the same as the probability of a flight time within any other 1-minute interval contained in the larger interval from 120 to 140 minutes.
- With every 1-minute interval being equally likely, the random variable x is said to have a uniform probability distribution.

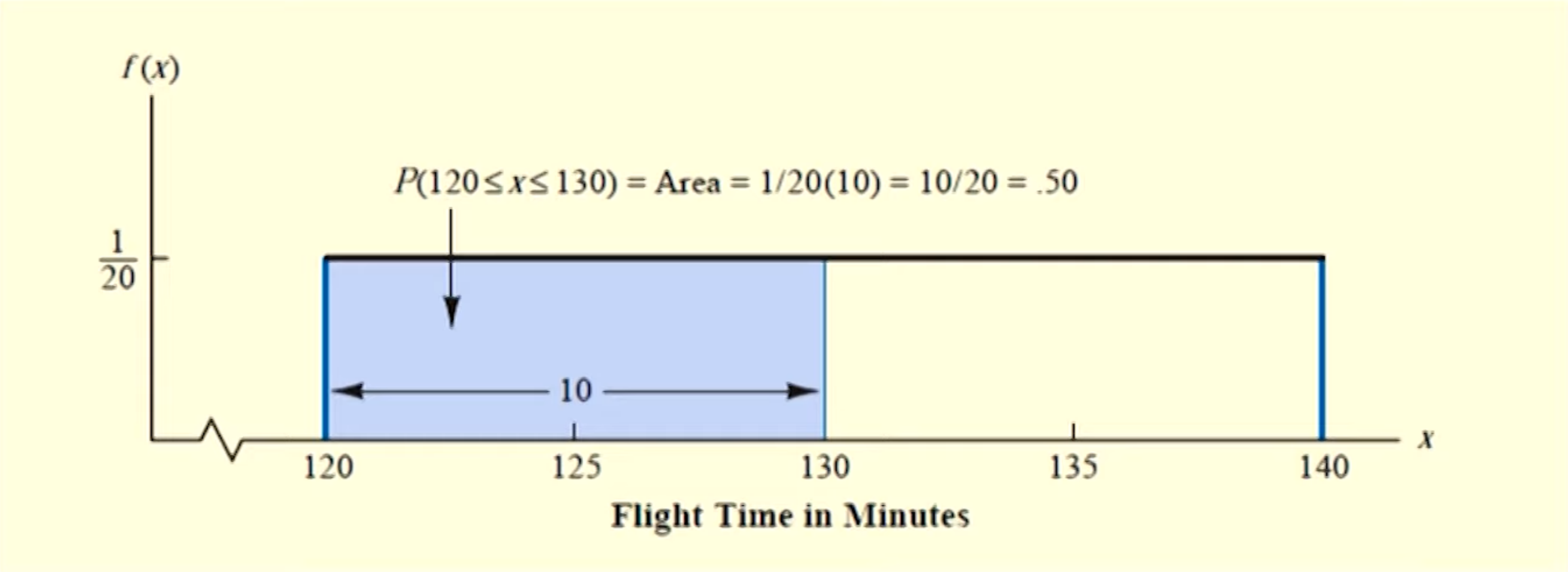
The Uniform Distribution- Example

Uniform Probability Distribution for Flight time



The Uniform Distribution- Example

Probability of a flight time between 120 and 130 minutes



Exponential Probability Distribution

- The exponential probability distribution is useful in describing the time it takes to complete a task.
- The exponential random variables can be used to describe:

Time between vehicle arrivals at a toll booth

Time required to complete a questionnaire

Distance between major defects in a highway

Exponential Probability Distribution

- Density Function

$$f(x) = \frac{1}{\mu} e^{-x/\mu}$$

where: μ = mean

$e = 2.71828$

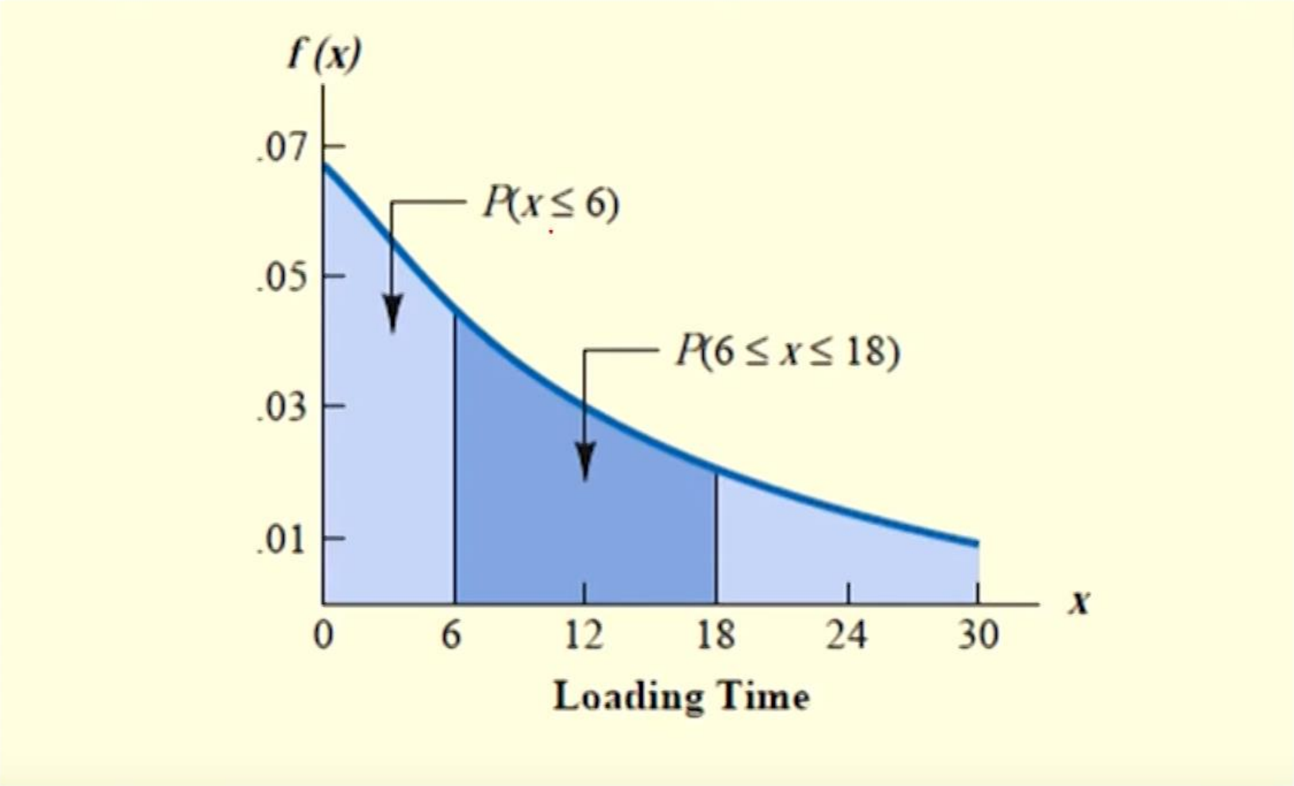
Exponential Probability Distribution

- Suppose that x represents the loading time for a truck at loading dock and follows such a distribution.
- If the mean, or average, loading time is 15 minutes (μ = 15), the appropriate probability density function for x is

$$f(x) = \frac{1}{15} e^{-x/15}$$

Exponential Probability Distribution

Exponential Distribution for the loading Dock Example



Exponential Probability Distribution

Cumulative Probabilities

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

where:

x0 = some specific value of x

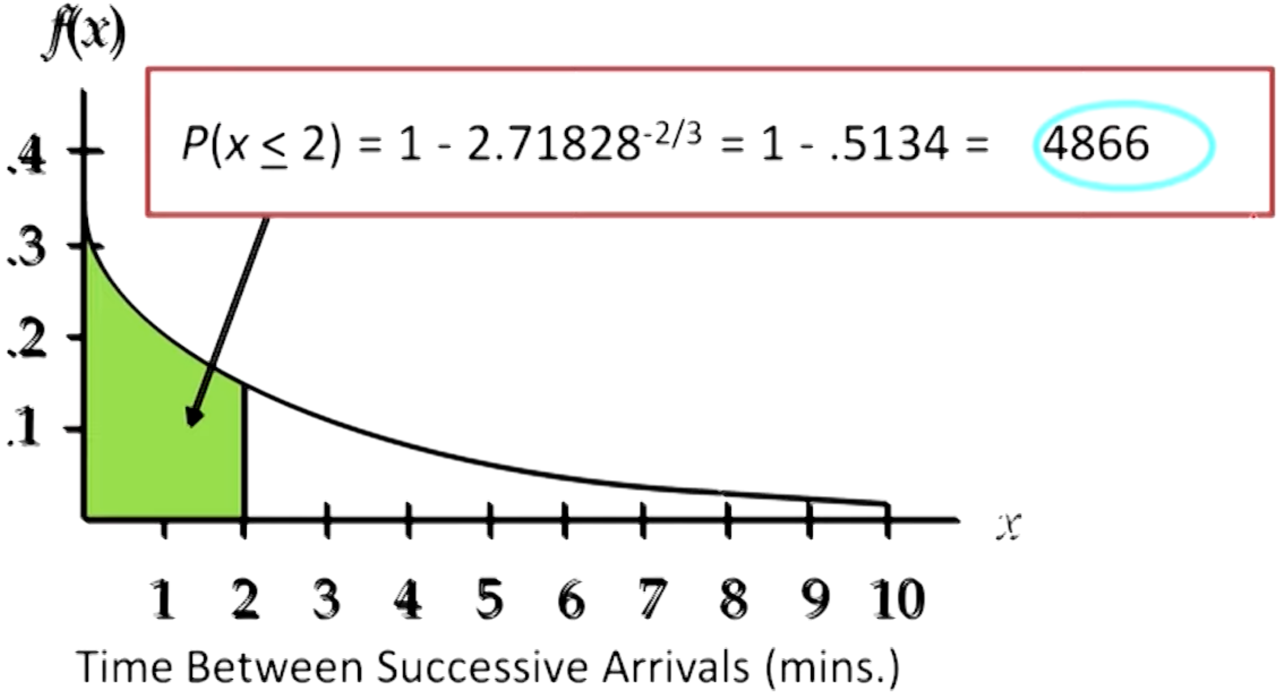
EduBridge

Exponential Probability Distribution- Ex

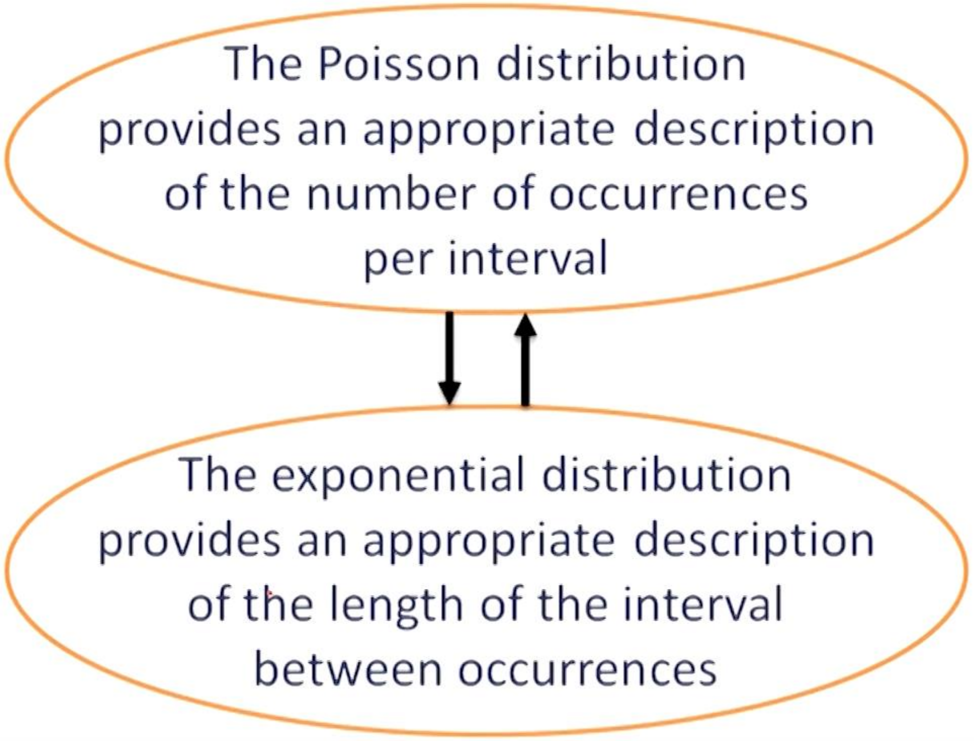
- The time between arrivals of cars at a Petrol pump follows an exponential probability distribution with a mean time between arrivals of 3 minutes.
- The Petrol pump owner would like to know the probability that the time between two successive arrivals will be 2 minutes or less.

Exponential Probability Distribution- Ex

Example: Petrol Pump Problem



Relationship b/w the Poisson and Exponential Distributions



Relationship b/w the Poisson and Exponential Distributions

Mean of Poisson and Mean of Exponential Distributions

- Because the average number of arrivals is 10 cars per hour, the average time between cars arriving is

$$\frac{1 \text{ hour}}{10 \text{ cars}} = .1 \text{ hour/car}$$

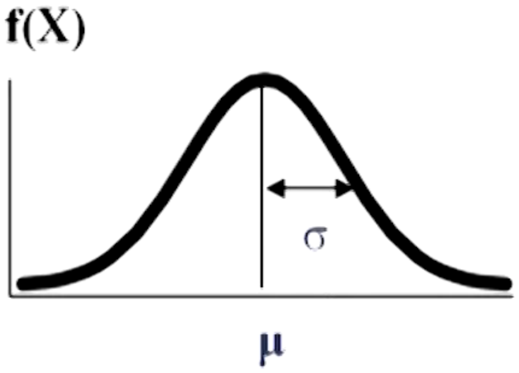
10

$\mu \rightarrow$ mean for Poisson

$\frac{1}{\mu} \rightarrow$ mean for Exp

The Normal Distribution: Properties

- 'Bell Shaped'
- Symmetrical
- Mean, Median and Mode are equal
- Location is characterized by the mean, μ
- Spread is characterized by the standard deviation, σ
- The random variable has an infinite theoretical range: $-\infty$ to $+\infty$



Mean = Median = Mode

The Normal Distribution: Density Function

The formula for the normal probability density function is

$$f(X) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{(X-\mu)}{\sigma}\right)^2}$$

Where e = the mathematical constant approximated by 2.71828

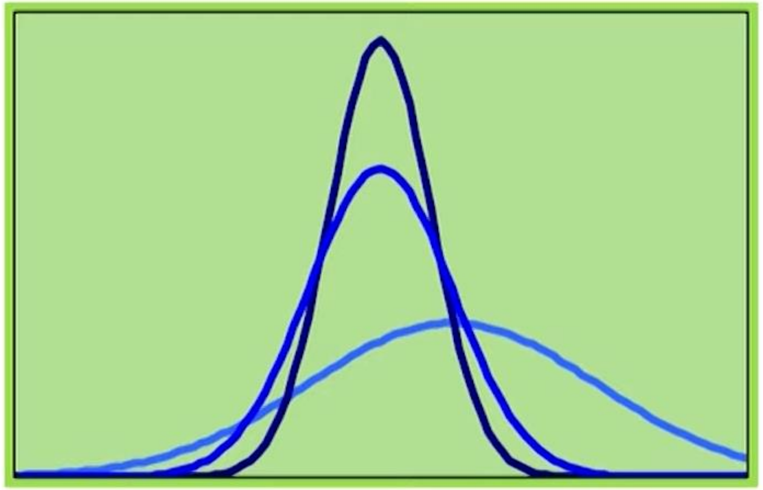
π = the mathematical constant approximated by 3.14159

μ = the population mean

σ = the population standard deviation

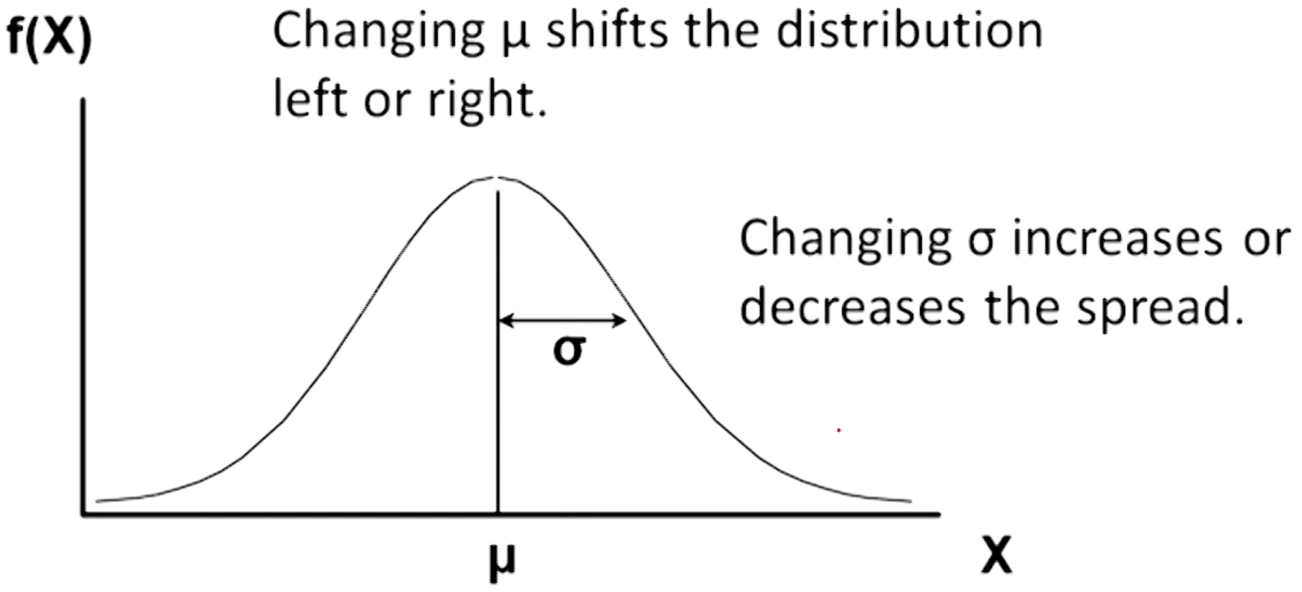
X = any value of the continuous variable

The Normal Distribution: Shape



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

The Normal Distribution: Shape



The Standardized Normal Distribution

- Any normal distribution (with any mean and standard deviation combination) can be transformed into the standardized normal distribution (Z).
- Need to transform X units into Z units.
- The standardized normal distribution has a mean of 0 and a standard deviation of 1.

The Standardized Normal Distribution

- Translate from X to the standardized normal (the "Z" distribution) by subtracting the mean of X and dividing by its standard deviation:

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

The Standardized Normal Distribution- DF

- The formula for the standardized normal probability density function is

$$f(Z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{Z^2}{2}}$$

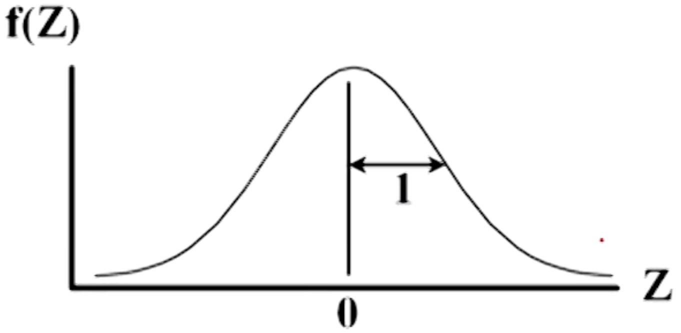
Where e = the mathematical constant approximated by 2.71828

π = the mathematical constant approximated by 3.14159

Z = any value of the standardized normal distribution

Standardized Normal Distribution-Shape

- Also known as the "Z" distribution
- Mean is 0
- Standard Deviation is 1



Values above the mean have positive Z-values, values below the mean have negative Z-values

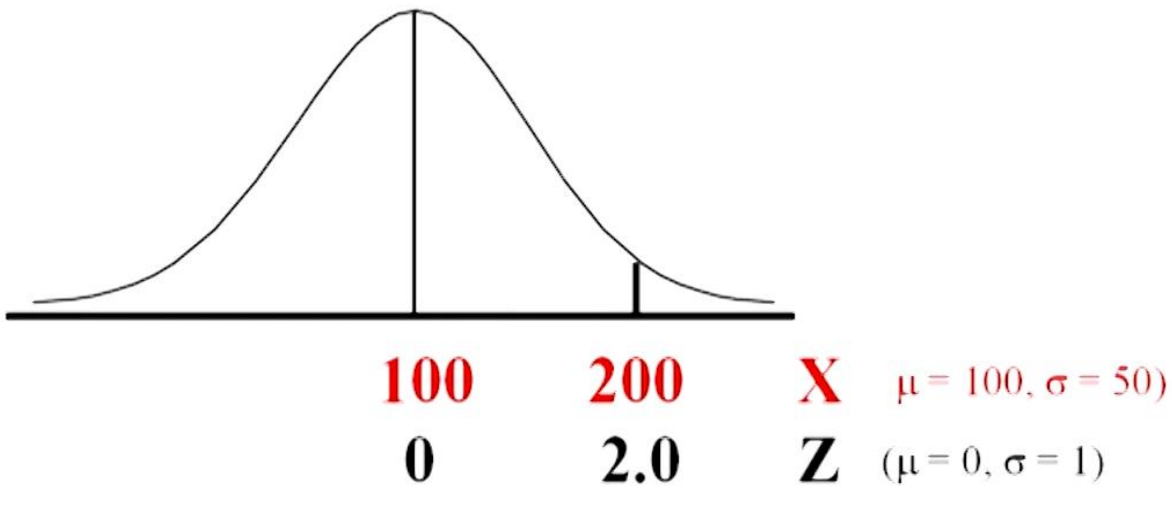
Standardized Normal Distribution-Ex

- If X is distributed normally with mean of 100 and standard deviation of 50, the Z value for $X = 200$ is

$$Z = \frac{X - \mu}{\sigma} = \frac{200 - 100}{50} = 2.0$$

- This says that $X = 200$ is two standard deviations (2 increments of 50 units) above the mean of 100.

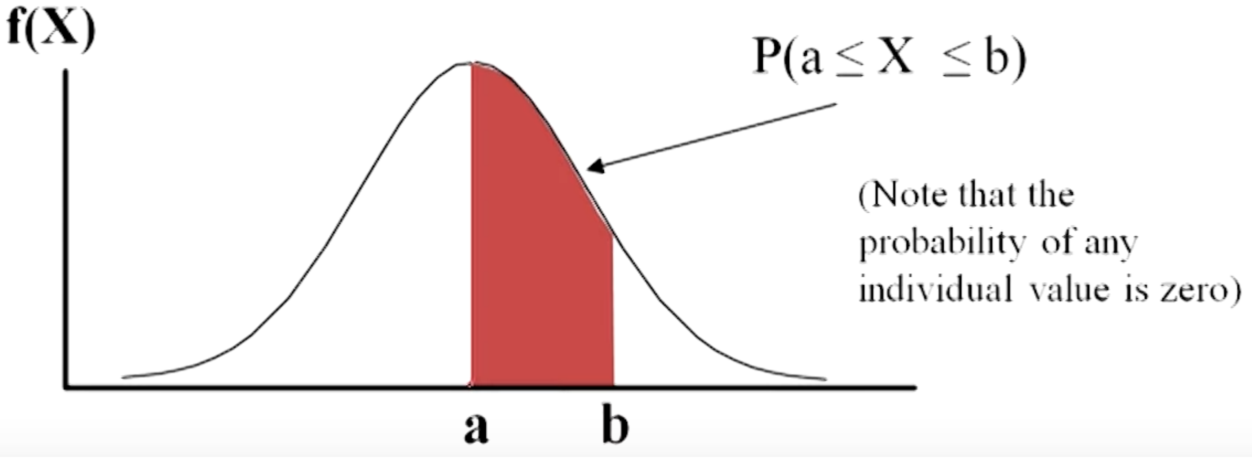
Standardized Normal Distribution-Ex



Note that the distribution is the same, only the scale has changed. We can express the problem in original units (X) or in standardized units (Z)

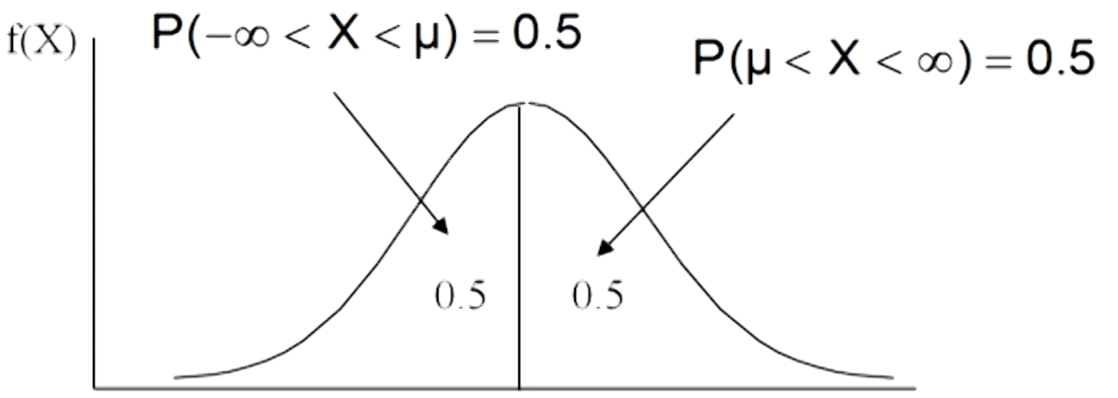
Standardized Normal Distribution-Ex

Probability is measured by the area under the curve



Standardized Normal Distribution-Ex

The total area under the curve is 1.0, and the curve is symmetric, so half is above the mean, half is below.



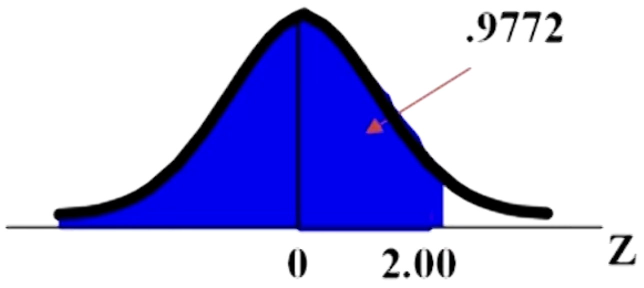
$$P(-\infty < X < \infty) = 1.0$$

Standardized Normal Distribution-Ex

Normal Probability Tables

Example:

$P(Z < 2.00) = .9772$

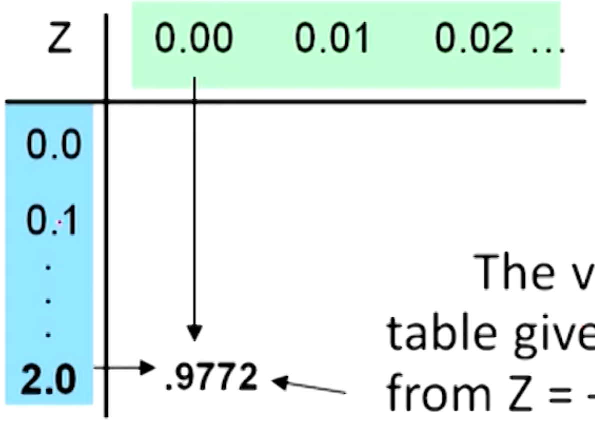


Standardized Normal Distribution-Ex

Normal Probability Tables

The column gives the value of Z to the second decimal point

The row shows the value of Z to the first decimal point



The value within the table gives the probability from Z = $-\infty$ up to the desired Z value.

P(Z < 2.00) = .9772

Standardized Normal Distribution-Ex

Finding Normal Probability Procedure

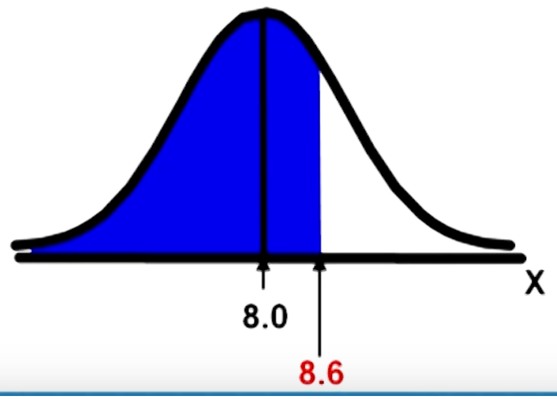
To find $P(a < X < b)$ when X is distributed normally:

- Draw the normal curve for the problem in terms of X.
- Translate X-values to Z-values.
- Use the Standardized Normal Table.

Standardized Normal Distribution-Ex

Finding Normal Probability: Example

- Let X represent the time it takes (in seconds) to download an image file from the internet.
- Suppose X is normal with mean 8.0 and standard deviation 5.0
- Find P(X < 8.6)

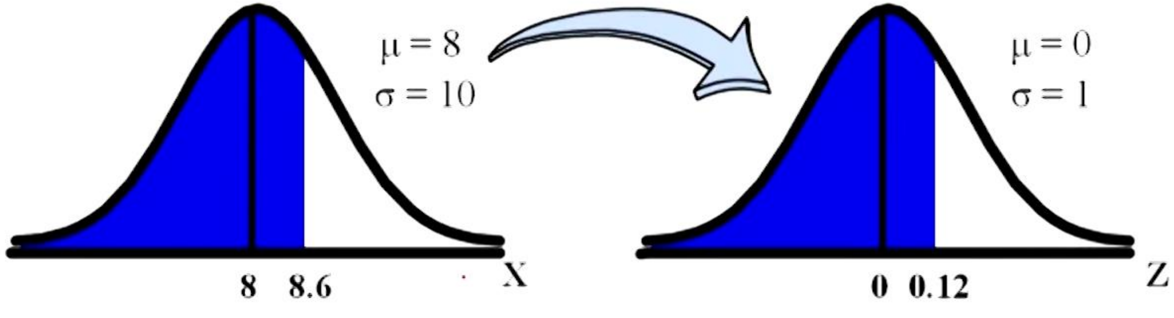


Standardized Normal Distribution-Ex

Finding Normal Probability: Example

- Suppose X is normal with mean 8.0 and standard deviation 5.0. Find P(X < 8.6).

$$Z = \frac{X-\mu}{\sigma} = \frac{8.6-8.0}{5.0} = 0.12$$



$P(X < 8.6)$ $P(Z < 0.12)$

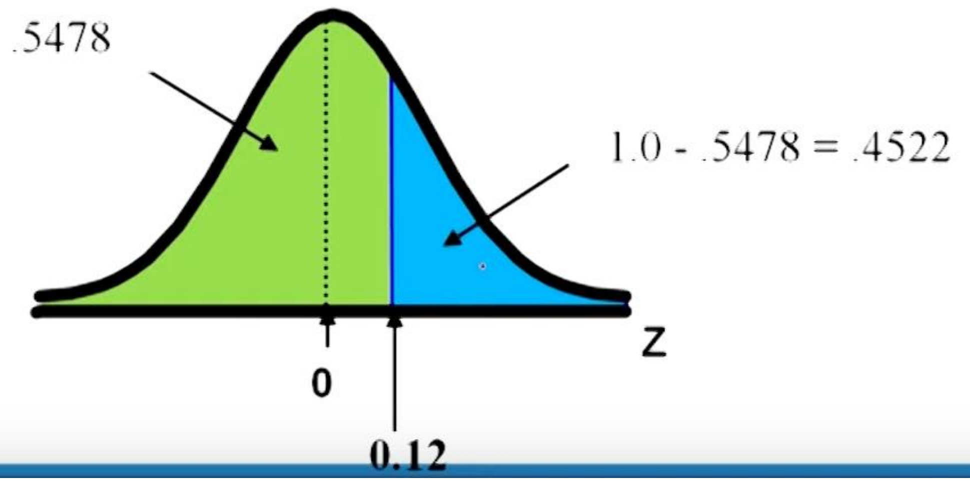
Standardized Normal Distribution-Ex

Finding Normal Probability: Example

- Find P(X > 8.6)…

$$P(X > 8.6) = P(Z > 0.12) = 1.0 - P(Z \leq 0.12)$$

$$= 1.0 - .5478 = .4522$$



Standardized Normal Distribution-Ex

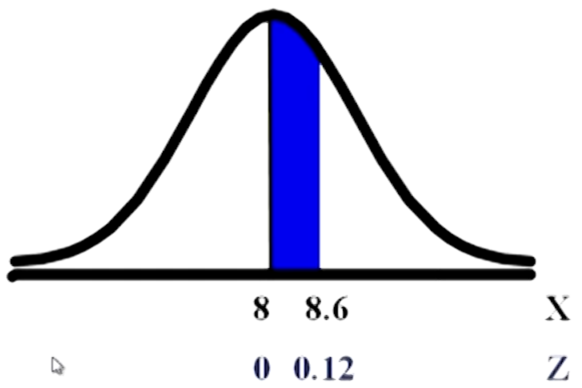
Finding Normal Probability: Between Two Values

- Suppose X is normal with mean 8.0 and standard deviation 5.0. Find P(8 < X < 8.6)

Calculate Z-values:

$$Z=\frac{X-\mu}{\sigma}=\frac{8-8}{5}=0$$

$$Z=\frac{X-\mu}{\sigma}=\frac{8.6-8}{5}=0.12$$



$P(8 < X < 8.6)$

$= P(0 < Z < 0.12)$

Standardized Normal Distribution-Ex

Finding Normal Probability Between Two Values

- Standardized Normal Probability
- Table (Portion)

Z	.00	.01	.02
0.0	.5000	.5040	.5080
0.1	.5398	.5438	.5478
0.2	.5793	.5832	.5871
0.3	.6179	.6217	.6255

$$P(8 < X < 8.6)$$
$$= P(0 < Z < 0.12)$$
$$= P(Z < 0.12) - P(Z \le 0)$$
$$= .5478 - .5000 = .0478$$