Math 183
Statistical Methods

Eddie Aamari
S.E.W. Assistant Professor

eaamari@ucsd.edu
math.ucsd.edu/~eaamari/
AP&M 5880A

Today: Review

- Develop good notational habits
- Expand the types of problems you have seen
- Enrich your problem-solving skill set

Probability Model

Is the numerical idea being

- Discrete? (i.e. 0,1,2,...)
  $\rightarrow$ Geometric, Binomial, Poisson, Negative Binomial
- Continuous?
  $\rightarrow$ Uniform, Exponential, Normal

Ask yourself this high-level question to help decide which category, then break out specific knowledge of each distribution.

Which model best summarizes the below idea?

The amount of time before the next volcano eruption on Earth

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The amount of time before the next volcano eruption on Earth

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 6. (Exponential)

Which model best summarizes the below idea?

The number of emails you will get today.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The number of emails you will get today.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 3. (Poisson)

Which model best summarizes the below idea?

The number of days we must wait until a politician next tells a lie.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The number of days we must wait until a politician next tells a lie.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 1. (Geometric)

Which model best summarizes the below idea?

The time of day that a meteor enters Earth's atmosphere.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The time of day that a meteor enters Earth's atmosphere.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 5. (Uniform)

Which model best summarizes the below idea?

The number of times an archer hits a target when shooting a quiver with ten arrows.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The number of times an archer hits a target when shooting a quiver with ten arrows.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 2. (Binomial)

Which model best summarizes the below idea?

The number of chocolates you must eat from a huge box of chocolates if you want to discover 4 with fruit fillings

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The number of chocolates you must eat from a huge box of chocolates if you want to discover 4 with fruit fillings

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 4. (Negative Binomial)

Which model best summarizes the below idea?

The length of people's noses

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The length of people's noses

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 7. (Normal)

Which model best summarizes the below idea?

The number of typos E. Aamari does on one of his lecture slides.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The number of typos E. Aamari does on one of his lecture slides.

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 3. (Poisson)

Which model best summarizes the below idea?

The number of people the TSA must screen before finding someone who forgot to dispose of his/her water bottle

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

The number of people the TSA must screen before finding someone who forgot to dispose of his/her water bottle

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 1. (Geometric)

Which model best summarizes the below idea?

The average IQ of the three contestants each day on the game jeopardy

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Which model best summarizes the below idea?

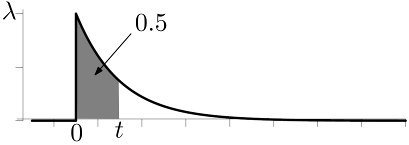
The average IQ of the three contestants each day on the game jeopardy

1. Geometric
2. Binomial
3. Poisson
4. Negative Binomial
5. Uniform
6. Exponential
7. Normal

Answer: 7. (Normal)

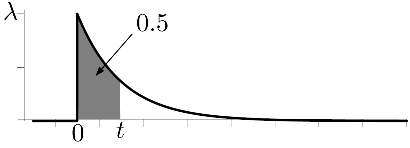
The median of a continuous random variable is the place on the x axis, say t, where 0.5 area is to the right of t under the density function.
Find the median of the exponential distribution.

The median of a continuous random variable is the place on the x axis, say t, where 0.5 area is to the right of t under the density function.
Find the median of the exponential distribution.



The median of a continuous random variable is the place on the x axis, say t, where 0.5 area is to the right of t under the density function.

Find the median of the exponential distribution.



We want $\int_0^t \lambda e^{-\lambda x} dx = 0.5$.

Thus, $-e^{-\lambda x}\Big|_0^t = 0.5$.

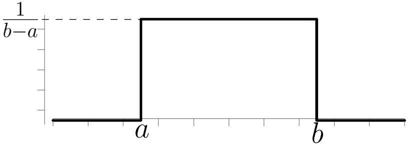
That is, $-e^{-\lambda t} + 1 = 0.5$.

Leading to $0.5 = e^{-\lambda t}$

So, $t = \dfrac{\ln 2}{\lambda}$.

You have seen that 68% of a Normal model is within 1 SD of the mean.
What percent of a uniform distribution is within 1 SD of the mean?

You have seen that 68% of a Normal model is within 1 SD of the mean.
What percent of a uniform distribution is within 1 SD of the mean?

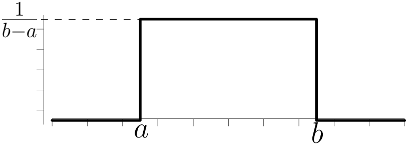


Recall that is $X = Unif(a, b)$,

$$E(X) = \frac{a+b}{2}$$

$$SD(X) = \frac{b-a}{\sqrt{12}}.$$

You have seen that 68% of a Normal model is within 1 SD of the mean.
What percent of a uniform distribution is within 1 SD of the mean?



Recall that is $X = Unif(a, b)$,

$$E(X) = \frac{a+b}{2}$$

$$SD(X) = \frac{b-a}{\sqrt{12}}.$$

If we go 1 SD to each side of the mean, this makes a rectangle with base $\frac{2(b-a)}{\sqrt{12}}$ and height $\frac{1}{b-a}$.

This product is $\frac{2}{\sqrt{12}} \simeq 0.577.$

Suppose that egg weights are normally distributed with a mean of 2 ounces and a standard deviation of 0.2 ounces. A carton of eggs at the store has 12 eggs.
If your goal is to find a carton with a total weight above 25 ounces, how many cartons should you expect to go through before finding one that is heavy enough?
What assumptions did you make in your solution?

Suppose that egg weights are normally distributed with a mean of 2 ounces and a standard deviation of 0.2 ounces. A carton of eggs at the store has 12 eggs.
If your goal is to find a carton with a total weight above 25 ounces, how many cartons should you expect to go through before finding one that is heavy enough?
What assumptions did you make in your solution?

Individual egg weights follow the random variable $W = N(2, 0.2)$. Thus, cartons are modeled by $C = W_1 + \ldots + W_{12}$.

Suppose that egg weights are normally distributed with a mean of 2 ounces and a standard deviation of 0.2 ounces. A carton of eggs at the store has 12 eggs.
If your goal is to find a carton with a total weight above 25 ounces, how many cartons should you expect to go through before finding one that is heavy enough?
What assumptions did you make in your solution?

Individual egg weights follow the random variable $W = N(2, 0.2)$. Thus, cartons are modeled by $C = W_1 + \ldots + W_{12}$.

Assuming that egg weights in a carton are independent, you get

$$\frac{C}{12} = \bar{W}_{12} = \frac{W_1 + \ldots + W_{12}}{12} = N\left(2, \frac{0.2}{\sqrt{12}}\right).$$

Thus, $C = N(24, \sqrt{12} \times 0.2) = N(24, 0.6928)$.

	Second decimal place					
Z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394
1.6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Since $C = N(24, 0.6928)$, writing $Z = \dfrac{C-24}{0.6928}$, you get

$$
\begin{aligned}
P(C \geq 25) &= P\left(Z \geq \frac{25-24}{0.6928}\right) \\
&\simeq P(Z \geq 1.44) \\
&= 1 - P(Z \leq 1.44). \\
&\simeq 1 - 0.9251 \\
&= 0.0759.
\end{aligned}
$$

	Second decimal place					
Z	0.00	0.01	0.02	0.03	0.04	0.05
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394
1.6	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

Since $C = N(24, 0.6928)$, writing $Z = \dfrac{C - 24}{0.6928}$, you get

$$
\begin{aligned}
P(C \geq 25) &= P\left(Z \geq \frac{25-24}{0.6928}\right) \\
&\simeq P(Z \geq 1.44) \\
&= 1 - P(Z \leq 1.44). \\
&\simeq 1 - 0.9251 \\
&= 0.0759.
\end{aligned}
$$

Let X be the number of carton we examine to find one that is heavy enough.
Then $X = Geom(p = 0.0749)$, and

$$E(X) = \frac{1}{p} \simeq 13.35.$$

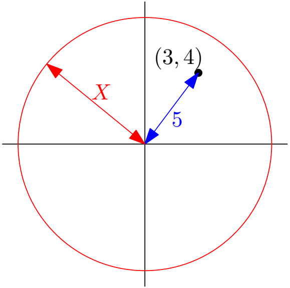
You would expect to go through 13.3 cartons before finding one that is heavy enough.

A random value following the exponential distribution $X = Exp(\lambda = 1/2)$ is generated. This value becomes the radius of a circle that is centered at the origin.

What is the probability that the point $(3, 4)$ will be inside the circle?

A random value following the exponential distribution $X = Exp(\lambda = 1/2)$ is generated. This value becomes the radius of a circle that is centered at the origin.

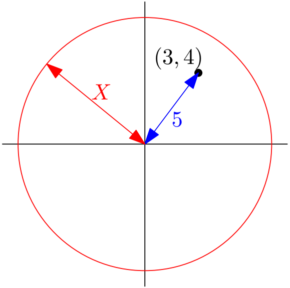
What is the probability that the point $(3, 4)$ will be inside the circle?



The circle contains the point $(3, 4)$ if $X > \sqrt{3^2 + 4^2} = 5$.

A random value following the exponential distribution $X = Exp(\lambda = 1/2)$ is generated. This value becomes the radius of a circle that is centered at the origin.

What is the probability that the point $(3, 4)$ will be inside the circle?



The circle contains the point $(3, 4)$ if $X > \sqrt{3^2 + 4^2} = 5$.

We want

$$
\begin{aligned}
P(X > 5) &= \int_5^\infty \frac{1}{2} e^{-x/2} dx \\
&= \lim_{n\to\infty} -e^{-x/2}\Big|_5^\infty \\
&= -\lim_{n\to\infty} \left(e^{-n/2} - e^{-5/2}\right) \\
&= e^{-5/2} \\
&=\simeq 0.082.
\end{aligned}
$$

Someone states that the expected value for the **area** of the circle should be $\pi \times E(X) \times E(X)$. Explain why this is incorrect and find the correct value.

Someone states that the expected value for the **area** of the circle should be $\pi \times E(X) \times E(X)$. Explain why this is incorrect and find the correct value.

The correct answer is $E(\pi X^2) = \pi \times E(X^2) \neq \pi \times E(X) \times E(X)$. Indeed, if $E(X^2)$ and $E(X) \times E(X)$ were equal, then the variance $Var(X) = E(X^2) - E(X)^2$ would be equal to zero.

Someone states that the expected value for the **area** of the circle should be $\pi \times E(X) \times E(X)$. Explain why this is incorrect and find the correct value.

The correct answer is $E(\pi X^2) = \pi \times E(X^2) \neq \pi \times E(X) \times E(X)$. Indeed, if $E(X^2)$ and $E(X) \times E(X)$ were equal, then the variance $Var(X) = E(X^2) - E(X)^2$ would be equal to zero.

To find $E(X^2)$, note that $Var(X) = \dfrac{1}{\lambda^2}$, and $E(X) = \dfrac{1}{\lambda}$.
Thus,

$$E(X^2) = Var(X) + E(X)^2 = \frac{1}{\lambda^2} + \left(\frac{1}{\lambda}\right)^2 = \frac{2}{\lambda^2} = \frac{2}{0.5^2} = 8.$$

Hence,

$$E(\pi X^2) = \pi E(X^2) = 8\pi.$$