Probability

Confuses us. But really neat. At times, continuous. At others, discrete.

Sample Space: \( \Omega \), \( \Pr[\omega] \).

Event: \( \Pr[A] = \sum_{\omega \in A} \Pr[\omega] \).

Random variables: \( X(\omega) \).

Distribution: \( \Pr[X=x] = \sum_{x} \Pr[X=x] = 1. \)

Random Variable: \( X \)

Event: \( A = [a, b] \), \( \Pr[X \in A] \).

CDF: \( F(x) = \Pr[X \leq x] \).

PDF: \( f(x) = \frac{dF(x)}{dx} \).

\( \int_{-\infty}^{\infty} f(x) \, dx = 1. \)
Probability

Probability!
Probability

Probability!
Confuses us.
Probability!
Confuses us. But really neat.
Probability!
Confuses us. But really neat.
At times,
Probability

Probability! Confuses us. But really neat. At times, continuous.
Probability!
Confuses us. But really neat.
At times, continuous. At others,
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.
Probability

Probability! Confuses us. But really neat. At times, continuous. At others, discrete.
Probability

Probability! Confuses us. But really neat. At times, continuous. At others, discrete.
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.
Probability

Probability! Confuses us. But really neat. At times, continuous. At others, discrete.
Probability

Probability! Confuses us. But really neat. At times, continuous. At others, discrete.
Probability

Probability! Confuses us. But really neat. At times, continuous. At others, discrete.
Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.
Probability

Probability! Confuses us. But really neat. At times, continuous. At others, discrete.
Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.
Sample Space: $\Omega$, $Pr[\omega]$. 
Probability

Confuses us. But really neat.
At times, continuous. At others, discrete.

Sample Space: $\Omega$, $Pr[\omega]$.
Event: $Pr[A] = \sum_{\omega \in A} Pr[\omega]$
$\sum_{\omega} Pr[\omega] = 1$. 
Probability

Probability! Confuses us. But really neat. At times, continuous. At others, discrete.

Sample Space: $\Omega$, $Pr[\omega]$. Event: $Pr[A] = \sum_{\omega \in A} Pr[\omega]$.

Random Variable: $X$

Event: $A = [a, b], Pr[X \in A]$.
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.

Sample Space: $\Omega$, $Pr[\omega]$.
Event: $Pr[A] = \sum_{\omega \in A} Pr[\omega]$ 
$\sum_{\omega} Pr[\omega] = 1$.
Random variables: $X(\omega)$.
Distribution: $Pr[X = x]$ 
$\sum_{x} Pr[X = x] = 1$.

Random Variable: $X$
Event: $A = [a, b], Pr[X \in A]$, 
CDF: $F(x) = Pr[X \leq x]$. 
PDF: $f(x) = \frac{dF(x)}{dx}$.
$\int_{-\infty}^{\infty} f(x) = 1$. 
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.

Sample Space: $\Omega$, $Pr[\omega]$.
Event: $Pr[A] = \sum_{\omega \in A} Pr[\omega]$
$\sum_{\omega} Pr[\omega] = 1$.
Random variables: $X(\omega)$.
Distribution: $Pr[X = x]$
$\sum_{x} Pr[X = x] = 1$.

Random Variable: $X$
Event: $A = [a, b]$, $Pr[X \in A]$
CDF: $F(x) = Pr[X \leq x]$.
PDF: $f(x) = \frac{dF(x)}{dx}$.
$\int_{-\infty}^{\infty} f(x) = 1$. 
Probability

Probability!
Confuses us. But really neat.
At times, continuous. At others, discrete.

Sample Space: $\Omega$, $\Pr[\omega]$.  
Event: $\Pr[A] = \sum_{\omega \in A} \Pr[\omega]$  
$\sum_{\omega} \Pr[\omega] = 1$.

Random variables: $X(\omega)$.  
Distribution: $\Pr[X = x]$  
$\sum_{x} \Pr[X = x] = 1$.

Continuous as Discrete.  
$\Pr[X \in [x, x + \delta]] \approx f(x)\delta$

Random Variable: $X$  
Event: $A = [a, b]$, $\Pr[X \in A]$,  
CDF: $F(x) = \Pr[X \leq x]$.  
PDF: $f(x) = \frac{dF(x)}{dx}$.  
$\int_{-\infty}^{\infty} f(x) = 1$. 

Probability Rules are all good.

Conditional Probability.
Probability Rules are all good.

Conditional Probability.

Events: $A$, $B$
Probability Rules are all good.

Conditional Probability.
Events: $A, B$
Probability Rules are all good.

Conditional Probability.
Events: $A, B$
Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$. 
Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$
Probability Rules are all good.

Conditional Probability.

Events: \( A, B \)

Discrete: “Heads”, “Tails”, \( X = 1 \), \( Y = 5 \).

Continuous: \( X \) in [0.2, 0.3]. \( X \in [0.2, 0.3] \) or \( X \in [0.4, 0.6] \).

Conditional Probability:

\[
Pr[A|B] = \frac{Pr[A \cap Pr[B]}{Pr[B]}
\]

\( Pr[\text{“Second Heads”}|\text{“First Heads”}] \),

\( Pr[X \in [.2, .3]|X \in [.2, .3] \) or \( X \in [.5, .6] \]).
Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$

$Pr[\text{“Second Heads”}|\text{“First Heads”}]$

$Pr[X \in [.2, .3]|X \in [.2, .3]$ or $X \in [.5, .6]]$.

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap B]$
Probability Rules are all good.

Conditional Probability.

Events: \( A, B \)

Discrete: “Heads”, “Tails”, \( X = 1, Y = 5 \).

Continuous: \( X \) in \([.2, .3]\). \( X \in [.2, .3] \) or \( X \in [.4, .6] \).

Conditional Probability: \( Pr[A|B] = \frac{Pr[A \cap Pr[B]}{Pr[B]} \)

\( Pr[ \text{“Second Heads”} | \text{“First Heads”} ] \),
\( Pr[X \in [.2, .3] | X \in [.2, .3] \text{ or } X \in [.5, .6] ] \).

Total Probability Rule: \( Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}] \)
\( Pr[ \text{“Second Heads”}] = Pr[HH] + Pr[HT] \)
Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$

$Pr[\text{“Second Heads”}|\text{“First Heads”}]$, $Pr[X \in [.2, .3]|X \in [.2, .3]$ or $X \in [.5, .6]]$.

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$

$Pr[\text{“Second Heads”}] = Pr[HH] + Pr[HT]$

$B$ is First coin heads.
Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap Pr[B]}{Pr[B]}$

$Pr["Second Heads"]|"First Heads"]$,

$Pr[X \in [.2, .3]|X \in [.2, .3] \text{ or } X \in [.5, .6]]$.

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$

$Pr["Second Heads"] = Pr[HH] + Pr[HT]$

$B$ is First coin heads.

$Pr[X \in [.45, .55]] = Pr[X \in [.45, .50]] + Pr[X \in (.50, .55]]$
Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap Pr[B]}{Pr[B]}$

$Pr[\text{“Second Heads”}|\text{“First Heads”}]$

$Pr[X \in [.2, .3]|X \in [.2, .3]$ or $X \in [.5, .6]]$.

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$

$Pr[\text{“Second Heads”}] = Pr[HH] + Pr[HT]$

$B$ is First coin heads.

$Pr[X \in [.45, .55]] = Pr[X \in [.45, .50]] + Pr[X \in (.50, .55]]$

$B$ is $X \in [0, .5]$.
Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in [.2, .3]. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$

$Pr[\text{“Second Heads”}|\text{“First Heads”}],$
$Pr[X \in [.2, .3]|X \in [.2, .3]$ or $X \in [.5, .6]]$.

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$

$Pr[\text{“Second Heads”}] = Pr[HH] + Pr[HT]$ $B$ is First coin heads.

$Pr[X \in [.45, .55]] = Pr[X \in [.45, .50]] + Pr[X \in (.50, .55]]$ $B$ is $X \in [0, .5]$. 
Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$

$Pr[\text{“Second Heads”}|\text{“First Heads”}],$

$Pr[X \in [.2, .3]|X \in [.2, .3]$ or $X \in [.5, .6]]$.

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \bar{B}]$

$Pr[\text{“Second Heads”}] = Pr[HH] + Pr[HT]$

$B$ is First coin heads.

$Pr[X \in [.45, .55]] = Pr[X \in [.45, .50]] + Pr[X \in (.50, .55]]$

$B$ is $X \in [0, .5]$


Probability Rules are all good.

Conditional Probability.

Events: $A, B$


Continuous: $X$ in $[.2, .3]$. $X \in [.2, .3]$ or $X \in [.4, .6]$.

Conditional Probability: $Pr[A|B] = \frac{Pr[A \cap B]}{Pr[B]}$

$Pr["Second Heads"]|"First Heads"]$,
$Pr[X \in [.2, .3]|X \in [.2, .3]$ or $X \in [.5, .6]]$.

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$

$Pr["Second Heads"] = Pr[HH] + Pr[HT]$

$B$ is First coin heads.

$Pr[X \in [.45, .55]] = Pr[X \in [.45, .50]] + Pr[X \in (.50, .55]]$

$B$ is $X \in [0, .5]$


All work for continuous with intervals as events.
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

The distribution of one of the variables.

\[ E[Y|X=1] = \left( \frac{.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5}{.44} \right) = 1.16 \]

\[ E[Y|X=2] = \left( \frac{.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5}{.32} \right) = 1.25 \]

\[ E[Y|X=4] = \left( \frac{.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5}{.18} \right) = 1.35 \]

\[ E[Y|X=8] = \left( \frac{.02 \times 1 + .02 \times 2 + .01 \times 3 + .01 \times 5}{.06} \right) = 1.10 \]

\[ E[Y] = E[E[Y|X]] = E[Pr[X = 3, Y = 3]] + E[Pr[X = 3, Y = 2]] + \cdots = 2.86 \]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Here is the marginal distribution for $X$:

- $P(X=1) = \frac{.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5}{.44} = .16$
- $P(X=2) = \frac{.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5}{.32} = .25$
- $P(X=4) = \frac{.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5}{.18} = .35$
- $P(X=8) = \frac{.02 \times 1 + .02 \times 2 + .01 \times 3 + .01 \times 5}{.06} = .10$

The conditional expectation of $Y$ given $X$ is calculated as follows:

- $E[Y|X=1] = \frac{.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5}{.44} = 1.16$
- $E[Y|X=2] = \frac{.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5}{.32} = 1.25$
- $E[Y|X=4] = \frac{.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5}{.18} = 1.35$
- $E[Y|X=8] = \frac{.02 \times 1 + .02 \times 2 + .01 \times 3 + .01 \times 5}{.06} = 1.10$
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
</tbody>
</table>

Marginal Distribution?

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
<td>.20</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
<td>.26</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
<td>.31</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
<td>.25</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
<td></td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one.
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another. The distribution of one of the variables.
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.

The distribution of one of the variables.

$E[Y|X]$?
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.

The distribution of one of the variables.

\[ E[Y|X] \]

\[ E[Y|X = 1] = (0.03 \times 1 + 0.2 \times 2 + 0.21 \times 3 + 0.02 \times 5)/0.44 = \frac{1.16}{44}. \]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another. The distribution of one of the variables.

\[ E[Y|X] \]

\[
E[Y|X = 1] = (0.03 \times 1 + 0.2 \times 2 + 0.21 \times 3 + 0.02 \times 5)/.44 = \frac{1.16}{44}.
\]

\[
E[Y|X = 2] = (0.05 \times 1 + 0.01 \times 2 + 0.06 \times 3 + 0.2 \times 5)/.32 = \frac{1.25}{32}.
\]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
<th>Pr[X = 3, Y = 3]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
<td>.20</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
<td>.26</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
<td>.31</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
<td>.25</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

\[ E[Y|X]? \]

\[ E[Y|X = 1] = (.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5)/.44 = \frac{1.16}{44}. \]
\[ E[Y|X = 2] = (.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5)/.32 = \frac{1.25}{32}. \]
\[ E[Y|X = 4] = (.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5)/.18 = \frac{.35}{18}. \]
## Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

**E[Y|X]?**

\[
E[Y|X = 1] = (0.03 \times 1 + 0.2 \times 2 + 0.21 \times 3 + 0.02 \times 5)/.44 = \frac{1.16}{44}.
\]

\[
E[Y|X = 2] = (0.05 \times 1 + 0.01 \times 2 + 0.06 \times 3 + 0.2 \times 5)/.32 = \frac{1.25}{32}.
\]

\[
E[Y|X = 4] = (0.1 \times 1 + 0.03 \times 2 + 0.03 \times 3 + 0.02 \times 5)/.18 = \frac{.35}{18}.
\]

\[
E[Y|X = 8] = (0.02 \times 1 + 0.02 \times 2 + 0.01 \times 3 + 0.01 \times 5)/.06 = \frac{.10}{06}.
\]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another. The distribution of one of the variables.

$E[Y|X]$?

$E[Y|X = 1] = (.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5)/.44 = \frac{1.16}{44}.$

$E[Y|X = 2] = (.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5)/.32 = \frac{1.25}{32}.$

$E[Y|X = 4] = (.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5)/.18 = \frac{.35}{18}.$

$E[Y|X = 8] = (.02 \times 1 + .02 \times 2 + .01 \times 3 + .01 \times 5)/.06 = \frac{.10}{.06}.$

$E[Y]$
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

\[
E[Y|X]\
\]

\[
E[Y|X = 1] = (.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5)/.44 = \frac{1.16}{44}.
\]

\[
E[Y|X = 2] = (.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5)/.32 = \frac{1.25}{32}.
\]

\[
E[Y|X = 4] = (.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5)/.18 = \frac{.35}{18}.
\]

\[
E[Y|X = 8] = (.02 \times 1 + .02 \times 2 + .01 \times 3 + .01 \times 5)/.06 = \frac{.10}{.06}.
\]

\[
E[Y] = E[E[Y|X]] =
\]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another. The distribution of one of the variables.

\[ E[Y|X] \]

\[
E[Y|X = 1] = (0.03 \times 1 + 0.2 \times 2 + 0.21 \times 3 + 0.02 \times 5)/0.44 = \frac{1.16}{44}. \\
E[Y|X = 2] = (0.05 \times 1 + 0.01 \times 2 + 0.06 \times 3 + 0.2 \times 5)/0.32 = \frac{1.25}{32}. \\
E[Y|X = 4] = (0.1 \times 1 + 0.03 \times 2 + 0.03 \times 3 + 0.02 \times 5)/0.18 = \frac{0.35}{18}. \\
E[Y|X = 8] = (0.02 \times 1 + 0.02 \times 2 + 0.01 \times 3 + 0.01 \times 5)/0.06 = \frac{0.10}{0.06}. \\
E[Y] = E[E[Y|X]] = E[Y|X = 1]Pr[X = 1] \]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another. The distribution of one of the variables.

\[ E[Y|X] \]

\[ E[Y|X = 1] = (0.03 \times 1 + 0.2 \times 2 + 0.21 \times 3 + 0.02 \times 5)/0.44 = \frac{1.16}{44} \] .
\[ E[Y|X = 2] = (0.05 \times 1 + 0.01 \times 2 + 0.06 \times 3 + 0.2 \times 5)/0.32 = \frac{1.25}{32} \] .
\[ E[Y|X = 4] = (0.1 \times 1 + 0.03 \times 2 + 0.03 \times 3 + 0.02 \times 5)/0.18 = \frac{0.35}{18} \] .
\[ E[Y|X = 8] = (0.02 \times 1 + 0.02 \times 2 + 0.01 \times 3 + 0.01 \times 5)/0.06 = \frac{0.10}{06} \] .

Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.

The distribution of one of the variables.

\[ E[Y|X]? \]

\[ E[Y|X = 1] = (0.03 \times 1 + 0.2 \times 2 + 0.21 \times 3 + 0.02 \times 5)/.44 = \frac{1.16}{.44} = 2.68. \]

\[ E[Y|X = 2] = (0.05 \times 1 + 0.01 \times 2 + 0.06 \times 3 + 0.2 \times 5)/.32 = \frac{1.25}{.32} = 3.97. \]

\[ E[Y|X = 4] = (0.1 \times 1 + 0.03 \times 2 + 0.03 \times 3 + 0.02 \times 5)/.18 = \frac{3.5}{.18} = 19.44. \]

\[ E[Y|X = 8] = (0.02 \times 1 + 0.02 \times 2 + 0.01 \times 3 + 0.01 \times 5)/.06 = 10.00. \]

Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
</tbody>
</table>

|       | .20 | .26 | .31 | .25 |

Marginal Distribution? Here is one. And here is another.

The distribution of one of the variables.

\[ E[Y|X]? \]

\[ E[Y|X = 1] = (.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5) / .44 = \frac{1.16}{44}. \]
\[ E[Y|X = 2] = (.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5) / .32 = \frac{1.25}{32}. \]
\[ E[Y|X = 4] = (.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5) / .18 = \frac{.35}{18}. \]
\[ E[Y|X = 8] = (.02 \times 1 + .02 \times 2 + .01 \times 3 + .01 \times 5) / .06 = \frac{.10}{.06}. \]


\[ E[Y] = (1.16 + 1.25 + .35 + .10) = 2.86. \]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
</tbody>
</table>

|   | .44 | .32 | .18 | .06 |

Marginal Distribution? Here is one. And here is another.

The distribution of one of the variables.

\[ E[Y|X] \]

\[ E[Y|X = 1] = (.03 \times 1 + .2 \times 2 + .21 \times 3 + .02 \times 5)/.44 = \frac{1.16}{44} \]
\[ E[Y|X = 2] = (.05 \times 1 + .01 \times 2 + .06 \times 3 + .2 \times 5)/.32 = \frac{1.25}{32} \]
\[ E[Y|X = 4] = (.1 \times 1 + .03 \times 2 + .03 \times 3 + .02 \times 5)/.18 = \frac{.35}{18} \]
\[ E[Y|X = 8] = (.02 \times 1 + .02 \times 2 + .01 \times 3 + .01 \times 5)/.06 = \frac{.10}{06} \]

\[ E[Y] = (1.16 + 1.25 + .35 + .10) = 2.86. \]
Joint distribution.

<table>
<thead>
<tr>
<th>Y/X</th>
<th>1</th>
<th>2</th>
<th>4</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>.03</td>
<td>.05</td>
<td>.1</td>
<td>.02</td>
</tr>
<tr>
<td>2</td>
<td>.2</td>
<td>.01</td>
<td>.03</td>
<td>.02</td>
</tr>
<tr>
<td>3</td>
<td>.21</td>
<td>.06</td>
<td>.03</td>
<td>.01</td>
</tr>
<tr>
<td>5</td>
<td>.02</td>
<td>.2</td>
<td>.02</td>
<td>.01</td>
</tr>
<tr>
<td></td>
<td>.44</td>
<td>.32</td>
<td>.18</td>
<td>.06</td>
</tr>
</tbody>
</table>

Marginal Distribution? Here is one. And here is another.

The distribution of one of the variables.

$E[Y|X]$?

$E[Y|X = 1] = (0.03 \times 1 + 0.2 \times 2 + 0.21 \times 3 + 0.02 \times 5)/0.44 = \frac{1.16}{44}$.

$E[Y|X = 2] = (0.05 \times 1 + 0.01 \times 2 + 0.06 \times 3 + 0.2 \times 5)/0.32 = \frac{1.25}{32}$.

$E[Y|X = 4] = (0.1 \times 1 + 0.03 \times 2 + 0.03 \times 3 + 0.02 \times 5)/0.18 = \frac{0.35}{18}$.

$E[Y|X = 8] = (0.02 \times 1 + 0.02 \times 2 + 0.01 \times 3 + 0.01 \times 5)/0.06 = \frac{0.10}{0.06}$.


$E[Y] = (1.16 + 1.25 + 0.35 + 0.10) = 2.86$. 
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where
\[
f_{X,Y}(x, y) = \frac{1}{|A|} \mathbb{1}\{(x, y) \in A\}
\]
where \(|A|\) is the area of \(A\).

Interpretation. Think of \((X, Y)\) as being discrete on a grid with mesh size \(\varepsilon\) and
\[
\Pr[X = m\varepsilon, Y = n\varepsilon] = f_{X,Y}(m\varepsilon, n\varepsilon)\varepsilon^2.
\]

Extension: \(X = (X_1, \ldots, X_n)\) with \(f_X(x)\).
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\).
One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y)\,dx\,dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y)\,dx\,dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y) \, dx \, dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X, Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X, Y}(x, y) \, dx\,dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X, Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

Example:
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y) \, dx \, dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

**Example:** Choose a point \((X, Y)\) uniformly in the set \(A \subset \mathbb{R}^2\).
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y)\,dx\,dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

**Example:** Choose a point \((X, Y)\) uniformly in the set \(A \subset \mathbb{R}^2\). Then

\[
f_{X,Y}(x, y) = \frac{1}{|A|} \mathbf{1}\{(x, y) \in A\}
\]
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y)\,dx\,dy = Pr[X \in (x, x + dx), \, Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

**Example:** Choose a point \((X, Y)\) uniformly in the set \(A \subset \mathbb{R}^2\). Then

\[
f_{X,Y}(x, y) = \frac{1}{|A|} 1\{(x, y) \in A\}
\]

where \(|A|\) is the area of \(A\).
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[ f_{X,Y}(x, y) \, dx \, dy = Pr[X \in (x, x + dx), Y \in (y + dy)]. \]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

**Example:** Choose a point \((X, Y)\) uniformly in the set \(A \subset \mathbb{R}^2\). Then

\[ f_{X,Y}(x, y) = \frac{1}{|A|} 1 \{ (x, y) \in A \} \]

where \(|A|\) is the area of \(A\).

**Interpretation.**
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y) \, dx \, dy = \Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

**Example:** Choose a point \((X, Y)\) uniformly in the set \(A \subset \mathbb{R}^2\). Then

\[
f_{X,Y}(x, y) = \frac{1}{|A|} 1\{(x, y) \in A\}
\]

where \(|A|\) is the area of \(A\).

**Interpretation.** Think of \((X, Y)\) as being discrete on a grid with mesh size \(\varepsilon\).
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y)\,dx\,dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

**Example:** Choose a point \((X, Y)\) uniformly in the set \(A \subset \mathbb{R}^2\). Then

\[
f_{X,Y}(x, y) = \frac{1}{|A|} 1\{(x, y) \in A\}
\]

where \(|A|\) is the area of \(A\).

**Interpretation.** Think of \((X, Y)\) as being discrete on a grid with mesh size \(\varepsilon\) and \(Pr[X = m\varepsilon, Y = n\varepsilon] = f_{X,Y}(m\varepsilon, n\varepsilon)\varepsilon^2\).
Multiple Continuous Random Variables

One defines a pair \((X, Y)\) of continuous RVs by specifying \(f_{X,Y}(x, y)\) for \(x, y \in \mathbb{R}\) where

\[
f_{X,Y}(x, y) \, dx\,dy = Pr[X \in (x, x + dx), Y \in (y + dy)].
\]

The function \(f_{X,Y}(x, y)\) is called the joint pdf of \(X\) and \(Y\).

**Example:** Choose a point \((X, Y)\) uniformly in the set \(A \subset \mathbb{R}^2\). Then

\[
f_{X,Y}(x, y) = \frac{1}{|A|} 1\{(x, y) \in A\}
\]

where \(|A|\) is the area of \(A\).

**Interpretation.** Think of \((X, Y)\) as being discrete on a grid with mesh size \(\varepsilon\) and \(Pr[X = m\varepsilon, Y = n\varepsilon] = f_{X,Y}(m\varepsilon, n\varepsilon)\varepsilon^2\).

**Extension:**
Multiple Continuous Random Variables

One defines a pair $(X, Y)$ of continuous RVs by specifying $f_{X,Y}(x, y)$ for $x, y \in \mathbb{R}$ where

$$f_{X,Y}(x, y)dx\,dy = \Pr[X \in (x, x + dx), Y \in (y + dy)].$$

The function $f_{X,Y}(x, y)$ is called the joint pdf of $X$ and $Y$.

**Example:** Choose a point $(X, Y)$ uniformly in the set $A \subset \mathbb{R}^2$. Then

$$f_{X,Y}(x, y) = \frac{1}{|A|} 1\{(x, y) \in A\}$$

where $|A|$ is the area of $A$.

**Interpretation.** Think of $(X, Y)$ as being discrete on a grid with mesh size $\varepsilon$ and $\Pr[X = m\varepsilon, Y = n\varepsilon] = f_{X,Y}(m\varepsilon, n\varepsilon)\varepsilon^2$.

**Extension:** $X = (X_1, \ldots, X_n)$ with $f_X(x)$. 
Example of Continuous $(X, Y)$
Example of Continuous $(X, Y)$

Pick a point $(X, Y)$ uniformly in the unit circle.
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X, Y}(x, y) = \frac{1}{\pi} \mathbb{1}_{x^2 + y^2 \leq 1}.
\]
Example of Continuous $(X, Y)$

Pick a point $(X, Y)$ uniformly in the unit circle.

\[ f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}. \]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
\implies f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Some events!

\[
Pr[X > 0, Y > 0] = 
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
\Rightarrow f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Some events:

\[
Pr[X > 0, Y > 0] = \frac{1}{4}
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[ f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}. \]

Some events!

\[
Pr[X > 0, Y > 0] = \frac{1}{4} \\
Pr[X < 0, Y > 0] =
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
\Rightarrow f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Some events!

\[
Pr[X > 0, Y > 0] = \frac{1}{4}
\]

\[
Pr[X < 0, Y > 0] = \frac{1}{4}
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
\implies f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Some events!

\[
Pr[X > 0, Y > 0] = \frac{1}{4}
\]
\[
Pr[X < 0, Y > 0] = \frac{1}{4}
\]
\[
Pr[X^2 + Y^2 \leq r^2] =
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
\implies f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Some events!

\[
Pr[X > 0, Y > 0] = \frac{1}{4}
\]

\[
Pr[X < 0, Y > 0] = \frac{1}{4}
\]

\[
Pr[X^2 + Y^2 \leq r^2] = r^2
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[ f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}. \]

Some events!

\[
\begin{align*}
Pr[X > 0, Y > 0] &= \frac{1}{4} \\
Pr[X < 0, Y > 0] &= \frac{1}{4} \\
Pr[X^2 + Y^2 \leq r^2] &= r^2 \\
Pr[X > Y] &=
\end{align*}
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
\implies f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Some events!

\[
\begin{align*}
Pr[X > 0, Y > 0] &= \frac{1}{4} \\
Pr[X < 0, Y > 0] &= \frac{1}{4} \\
Pr[X^2 + Y^2 \leq r^2] &= r^2 \\
Pr[X > Y] &= \frac{1}{2}.
\end{align*}
\]
Example of Continuous \((X, Y)\)
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X, Y}(x, y) = \frac{1}{\pi} \mathbb{1}_{x^2 + y^2 \leq 1}.
\]

Marginals:

\[
f_X(x) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) \, dy = \frac{2\pi}{\sqrt{1 - x^2}}.
\]

\[
f_Y(y) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) \, dx = \frac{2\pi}{\sqrt{1 - y^2}}.
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X, Y}(x, y) = \frac{1}{\pi} \mathbb{1}_{x^2 + y^2 \leq 1}.
\]

Marginals:

\[
f_X(x) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) \, dy = \frac{2}{\pi} \sqrt{1 - x^2},
\]

\[
f_Y(y) = \int_{-\infty}^{\infty} f_{X, Y}(x, y) \, dx = \frac{2}{\pi} \sqrt{1 - y^2}.
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x, y) = \frac{1}{\pi} \mathbb{1}\{x^2 + y^2 \leq 1\}.
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Marginals?
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Marginals?

\[
f_X(x) =
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Marginals?

\[
f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, dy
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x, y) = \frac{1}{\pi} \mathbf{1}\{x^2 + y^2 \leq 1\}.
\]

Marginals?

\[
f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, dy = \frac{2}{\pi} \sqrt{1 - x^2}
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Marginals?

\[
f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) \, dy = \frac{2}{\pi} \sqrt{1 - x^2}
\]

\[
f_Y(y) =
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Marginals?

\[
f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dy = \frac{2}{\pi} \sqrt{1-x^2}
\]

\[
f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y)\,dx
\]
Example of Continuous \((X, Y)\)

Pick a point \((X, Y)\) uniformly in the unit circle.

\[
f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.
\]

Marginals?

\[
f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dy = \frac{2}{\pi} \sqrt{1-x^2}
\]

\[
f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x,y) \, dx = \frac{2}{\pi} \sqrt{1-y^2}
\]
The exponential distribution with parameter $\lambda > 0$ is defined by

$$f_X(x) = \lambda e^{-\lambda x} \{x \geq 0\}$$

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - e^{-\lambda x}, & \text{if } x \geq 0. \end{cases}$$

Note that $\Pr[X > t] = e^{-\lambda t}$ for $t > 0$. 
Expo($\lambda$)

The exponential distribution with parameter $\lambda > 0$ is defined by

$$f_X(x) = \lambda e^{-\lambda x}1\{x \geq 0\}$$
The exponential distribution with parameter $\lambda > 0$ is defined by

$$f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$$

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - e^{-\lambda x}, & \text{if } x \geq 0. \end{cases}$$
The exponential distribution with parameter $\lambda > 0$ is defined by

$$f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$$

$$F_X(x) = \begin{cases} 0, & \text{if } x < 0 \\ 1 - e^{-\lambda x}, & \text{if } x \geq 0. \end{cases}$$

Note that $Pr[X > t] = e^{-\lambda t}$ for $t > 0$. 
Some Properties

1. Expo is memoryless. Let $X = \text{Expo}(\lambda)$. Then, for $s, t > 0$,\[
\Pr[X > t + s | X > s] = \Pr[X > t + s] = e^{-\lambda (t + s)} = e^{-\lambda t} = \Pr[X > t].
\]

2. Scaling Expo. Let $X = \text{Expo}(\lambda)$ and $Y = aX$ for some $a > 0$. Then\[
\Pr[Y > t] = \Pr[aX > t] = \Pr[X > t/a] = e^{-\lambda (t/a)} = e^{-\lambda (t/a)} = \Pr[Z > t]
\]
for $Z = \text{Expo}(\frac{\lambda}{a})$. Thus, $a \times \text{Expo}(\lambda) = \text{Expo}(\frac{\lambda}{a})$. Also, $\text{Expo}(\lambda) = \frac{\lambda}{1} \text{Expo}(1)$. 
Some Properties

1. *Expo* is memoryless.
Some Properties

1. *Expo is memoryless*. Let $X = Expo(\lambda)$.
Some Properties

1. *Expo is memoryless.* Let $X = Expo(\lambda)$. Then, for $s, t > 0$, 

$$
\Pr[X > t + s | X > s] = \Pr[X > t] = e^{-\lambda t}.
$$
Some Properties

1. *Expo is memoryless.* Let $X = Expo(\lambda)$. Then, for $s, t > 0$,

$$Pr[X > t + s \mid X > s] =$$
Some Properties

1. *Expo is memoryless.* Let $X = Expo(\lambda)$. Then, for $s, t > 0$,

$$Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}$$
Some Properties

1. **Expo is memoryless.** Let $X = Expo(\lambda)$. Then, for $s, t > 0$,

$$Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} =$$
Some Properties

1. **Expo is memoryless.** Let $X = Expo(\lambda)$. Then, for $s, t > 0,$

$$Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}$$
1. **Expo is memoryless.** Let \( X = Expo(\lambda) \). Then, for \( s, t > 0 \),

\[
Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}
\]

\[
= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}
\]

\[
= Pr[X > t].
\]
Some Properties

1. *Expo is memoryless.* Let \( X = Expo(\lambda) \). Then, for \( s, t > 0 \),

\[
Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}
\]

\[
= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}
\]

\[
= Pr[X > t].
\]

‘Used is a good as new.’
Some Properties

1. **Expo is memoryless.** Let $X = \text{Expo}(\lambda)$. Then, for $s, t > 0$,

   $$Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}$$

   $$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}$$

   $$= Pr[X > t].$$

   ‘Used is a good as new.’

2. **Scaling Expo.**
Some Properties

1. *Expo is memoryless.* Let $X = \text{Expo}(\lambda)$. Then, for $s, t > 0$,

$$
Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = Pr[X > t].
$$

‘Used is a good as new.’

2. **Scaling Expo.** Let $X = \text{Expo}(\lambda)$ and $Y = aX$ for some $a > 0$. 

```markdown
`e`
```
Some Properties

1. **Expo is memoryless.** Let $X = Expo(\lambda)$. Then, for $s, t > 0$,

$$Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}$$

$$= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}$$

$$= Pr[X > t].$$

‘Used is a good as new.’

2. **Scaling Expo.** Let $X = Expo(\lambda)$ and $Y = aX$ for some $a > 0$. Then

$$Pr[Y > t] =$$
Some Properties

1. **Expo is memoryless.** Let \( X = Expo(\lambda) \). Then, for \( s, t > 0 \),

\[
Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = Pr[X > t].
\]

‘Used is a good as new.’

2. **Scaling Expo.** Let \( X = Expo(\lambda) \) and \( Y = aX \) for some \( a > 0 \). Then

\[
Pr[Y > t] = Pr[aX > t] =
\]
Some Properties

1. **Expo is memoryless.** Let $X = Expo(\lambda)$. Then, for $s, t > 0,$

\[
Pr[X > t + s | X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}
\]

\[
= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}
\]

\[
= Pr[X > t].
\]

‘Used is a good as new.’

2. **Scaling Expo.** Let $X = Expo(\lambda)$ and $Y = aX$ for some $a > 0$. Then

\[
Pr[Y > t] = Pr[aX > t] = Pr[X > t/a]
\]
Some Properties

1. **Expo is memoryless.** Let \( X = \text{Expo}(\lambda) \). Then, for \( s, t > 0 \),

\[
\Pr[X > t + s \mid X > s] = \frac{\Pr[X > t + s]}{\Pr[X > s]}
= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}
= \Pr[X > t].
\]

‘Used is a good as new.’

2. **Scaling Expo.** Let \( X = \text{Expo}(\lambda) \) and \( Y = aX \) for some \( a > 0 \). Then

\[
\Pr[Y > t] = \Pr[aX > t] = \Pr[X > t/a]
= e^{-\lambda(t/a)} = e^{-(\lambda/a)t} =
\]
Some Properties

1. **Expo is memoryless.** Let \( X = Expo(\lambda) \). Then, for \( s, t > 0 \),

\[
Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}
= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}
= Pr[X > t].
\]

‘Used is a good as new.’

2. **Scaling Expo.** Let \( X = Expo(\lambda) \) and \( Y = aX \) for some \( a > 0 \). Then

\[
Pr[Y > t] = Pr[aX > t] = Pr[X > t/a]
= e^{-\lambda(t/a)} = e^{-(\lambda/a)t} = Pr[Z > t] \text{ for } Z = Expo(\lambda/a).
\]
1. **Expo is memoryless.** Let $X = \text{Expo}(\lambda)$. Then, for $s, t > 0,$

$$Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]} = \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t} = Pr[X > t].$$

‘Used is a good as new.’

2. **Scaling Expo.** Let $X = \text{Expo}(\lambda)$ and $Y = aX$ for some $a > 0$. Then

$$Pr[Y > t] = Pr[aX > t] = Pr[X > t/a] = e^{-\lambda(t/a)} = e^{-\frac{\lambda}{a}t} = Pr[Z > t] \text{ for } Z = \text{Expo}(\lambda/a).$$

Thus, $a \times \text{Expo}(\lambda) = \text{Expo}(\lambda/a)$. 

---

**Some Properties**
Some Properties

1. **Expo is memoryless.** Let $X = Expo(\lambda)$. Then, for $s, t > 0$,

\[
Pr[X > t + s \mid X > s] = \frac{Pr[X > t + s]}{Pr[X > s]}
= \frac{e^{-\lambda(t+s)}}{e^{-\lambda s}} = e^{-\lambda t}
= Pr[X > t].
\]

‘Used is a good as new.’

2. **Scaling Expo.** Let $X = Expo(\lambda)$ and $Y = aX$ for some $a > 0$. Then

\[
Pr[Y > t] = Pr[aX > t] = Pr[X > t/a]
= e^{-\lambda(t/a)} = e^{-(\lambda/a)t} = Pr[Z > t] \text{ for } Z = Expo(\lambda/a).
\]

Thus, $a \times Expo(\lambda) = Expo(\lambda/a)$.

Also, $Expo(\lambda) = \frac{1}{\lambda} Expo(1)$. 
More Properties

3. Scaling Uniform.

Let $X = U[0,1]$ and $Y = a + bX$ where $b > 0$.

Then, $\Pr[Y \in (y, y+\delta)] = \Pr[a+bX \in (y, y+\delta)] = \Pr[X \in (y-b, y+\delta-b)] = \frac{1}{b}\delta$, for $0 < y-b < 1 = \frac{1}{b}\delta$, for $a < y < a+b$.

Thus, $f_Y(y) = \frac{1}{b}$ for $a < y < a+b$.

Hence, $Y = U[a, a+b]$.

Replacing $b$ by $b-a$ we see that, if $X = U[0,1]$, then $Y = a + (b-a)X$ is $U[a, b]$. 
More Properties

3. Scaling Uniform.

Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$.

Then, $\Pr[Y \in (y, y + \delta)] = \Pr[a + bX \in (y, y + \delta)] = \Pr[X \in (y - ab, y + \delta - ab)] = \Pr[X \in (y - ab, y + \delta b)]$, for $0 < y - ab < 1 = \frac{1}{b}\delta$, for $a < y < a + b$.

Thus, $f_Y(y) = \frac{1}{b}$ for $a < y < a + b$.

Replacing $b$ by $b - a$ we see that, if $X = U[0, 1]$, then $Y = a + (b - a)X$ is $U[a, b]$. 

3. Scaling Uniform. Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. 
3. Scaling Uniform. Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] =$$
3. Scaling Uniform. Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y - a}{b}, \frac{y + \delta - a}{b}\right)]$$
3. Scaling Uniform. Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)$$
$$= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a+\delta}{b}\right)] =$$
3. **Scaling Uniform.** Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

\[
Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left( \frac{y-a}{b}, \frac{y+\delta-a}{b} \right)]
\]

\[
= Pr[X \in \left( \frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b} \right)] = \frac{1}{b} \delta, \text{ for}
\]
3. Scaling Uniform. Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]$$

$$= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] = \frac{1}{b} \delta, \text{ for } 0 < \frac{y-a}{b} < 1$$
3. **Scaling Uniform.** Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

$$
Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]
$$

$$
= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] = \frac{1}{b} \delta, \text{ for } 0 < \frac{y-a}{b} < 1
$$

$$
= \frac{1}{b} \delta, \text{ for } a < y < a+b.
$$
3. Scaling Uniform. Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]$$

$$= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] = \frac{1}{b} \delta, \text{ for } 0 < \frac{y-a}{b} < 1$$

$$= \frac{1}{b} \delta, \text{ for } a < y < a+b.$$ 

Thus, $f_Y(y) = \frac{1}{b}$ for $a < y < a+b$. 


3. Scaling Uniform. Let $X = U[0, 1]$ and $Y = a + bX$ where $b > 0$. Then,

$$Pr[ Y \in (y, y + \delta)] = Pr[ a + bX \in (y, y + \delta)] = Pr[ X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})]$$

$$= Pr[ X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = \frac{1}{b} \delta, \text{ for } 0 < \frac{y-a}{b} < 1$$

$$= \frac{1}{b} \delta, \text{ for } a < y < a + b.$$ 

Thus, $f_Y(y) = \frac{1}{b}$ for $a < y < a + b$. Hence, $Y = U[a, a+b]$. 

3. **Scaling Uniform.** Let \( X = U[0, 1] \) and \( Y = a + bX \) where \( b > 0 \). Then,

\[
\Pr[Y \in (y, y + \delta)] = \Pr[a + bX \in (y, y + \delta)] = \Pr[X \in \left( \frac{y-a}{b}, \frac{y+\delta-a}{b} \right)] = \Pr[X \in \left( \frac{y-a}{b}, \frac{y-a + \delta}{b} \right)] = \frac{1}{b} \delta, \text{ for } 0 < \frac{y-a}{b} < 1
\]

Thus, \( f_Y(y) = \frac{1}{b} \) for \( a < y < a + b \). Hence, \( Y = U[a, a+b] \).

Replacing \( b \) by \( b - a \) we see that, if \( X = U[0, 1] \), then \( Y = a + (b - a)X \) is \( U[a, b] \).
Some More Properties

4. Scaling pdf.

Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

$$\Pr[Y \in (y, y+\delta)] = \Pr[a + bX \in (y, y+\delta)] = \Pr[X \in (y-a/b, y+\delta-a/b)] = \Pr[X \in (y-a/b, y-a/b+\delta b)].$$

Now, the left-hand side is $f_Y(y)\delta$. Hence,

$$f_Y(y) = \frac{1}{b} f_X(y-a/b)\delta.$$
Some More Properties

4. Scaling pdf.

Let \( f_X(x) \) be the pdf of \( X \) and \( Y = a + bX \) where \( b > 0 \). Then

\[
Pr\left[ Y \in (y, y+\delta) \right] = Pr\left[ a + bX \in (y, y+\delta) \right] = Pr\left[ X \in (y-a/b, y+\delta-a/b) \right] = f_X(y-a/b) \delta \frac{1}{b}.
\]

Now, the left-hand side is \( f_Y(y) \delta \).

Hence, \( f_Y(y) = \frac{1}{b} f_X(y-a/b) \).
4. Scaling pdf. Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. 
4. **Scaling pdf.** Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] =$$
4. Scaling pdf. Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

$$
Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y - a}{b}, \frac{y + \delta - a}{b}\right)]
$$
4. **Scaling pdf.** Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]$$

$$= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] =$$
Some More Properties

4. Scaling pdf. Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]$$

$$= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] = f_X\left(\frac{y-a}{b}\right)\frac{\delta}{b}.$$
4. **Scaling pdf.** Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

$$Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]$$

$$= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] = f_X\left(\frac{y-a}{b}\right)\frac{\delta}{b}.$$ 

Now, the left-hand side is
4. **Scaling pdf.** Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

\[
Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]
\]

\[
= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] = f_X\left(\frac{y-a}{b}\right)\frac{\delta}{b}.
\]

Now, the left-hand side is $f_Y(y)\delta$. 
4. Scaling pdf. Let \( f_X(x) \) be the pdf of \( X \) and \( Y = a + bX \) where \( b > 0 \). Then

\[
Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left( \frac{y-a}{b}, \frac{y+\delta-a}{b} \right)] \\
= Pr[X \in \left( \frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b} \right)] = f_X\left( \frac{y-a}{b} \right) \frac{\delta}{b}.
\]

Now, the left-hand side is \( f_Y(y) \delta \). Hence,

\[
f_Y(y) = \frac{1}{b} f_X\left( \frac{y-a}{b} \right).
\]
**4. Scaling pdf.** Let $f_X(x)$ be the pdf of $X$ and $Y = a + bX$ where $b > 0$. Then

$$
Pr[Y \in (y, y + \delta)] = Pr[a + bX \in (y, y + \delta)] = Pr[X \in \left(\frac{y-a}{b}, \frac{y+\delta-a}{b}\right)]
$$

$$
= Pr[X \in \left(\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b}\right)] = f_X\left(\frac{y-a}{b}\right) \frac{\delta}{b}.
$$

Now, the left-hand side is $f_Y(y) \delta$. Hence,

$$
f_Y(y) = \frac{1}{b} f_X\left(\frac{y-a}{b}\right).
$$
Expectation

Definition:

The expectation of a random variable $X$ with pdf $f_X(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx.$$ 

Justification:

Say $X = n \delta$ w.p. $f_X(n \delta) \delta$ for $n \in \mathbb{Z}$.

Then,

$$E[X] = \sum_n n \delta \Pr[X = n \delta] = \sum_n n \delta f_X(n \delta) \delta = \int_{-\infty}^{\infty} x f_X(x) \, dx.$$ 

Indeed, for any $g$, one has

$$\int g(x) \, dx \approx \sum_n g(n \delta) \delta.$$ 

Choose $g(x) = xf_X(x)$.
Expectation

**Definition:** The expectation of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx.$$
Expectation

**Definition:** The *expectation* of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) dx.$$
**Expectation**

**Definition:** The *expectation* of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx.$$ 

**Justification:**
Expectation

**Definition:** The expectation of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) \, dx.$$ 

**Justification:** Say $X = n\delta$ w.p. $f_X(n\delta)\delta$ for $n \in \mathbb{Z}$. 

Indeed, for any $g$, one has

$$\int g(x) \, dx \approx \sum n g(n\delta) \delta.$$ 

Choose $g(x) = xf_X(x)$. 

Expectation

**Definition:** The expectation of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} xf_{X}(x)\,dx.$$  

**Justification:** Say $X = n\delta$ w.p. $f_{X}(n\delta)\delta$ for $n \in \mathbb{Z}$. Then,

$$E[X] = \sum_{n} (n\delta) Pr[X = n\delta]$$
**Expectation**

**Definition:** The *expectation* of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) \, dx.$$ 

**Justification:** Say $X = n\delta$ w.p. $f_X(n\delta)\delta$ for $n \in \mathbb{Z}$. Then,

$$E[X] = \sum_n (n\delta) \Pr[X = n\delta] = \sum_n (n\delta)f_X(n\delta)\delta$$
**Expectation**

**Definition:** The **expectation** of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) \, dx.$$  

**Justification:** Say $X = n\delta$ w.p. $f_X(n\delta)\delta$ for $n \in \mathbb{Z}$. Then,

$$E[X] = \sum_n (n\delta) \Pr[X = n\delta] = \sum_n (n\delta)f_X(n\delta)\delta = \int_{-\infty}^{\infty} xf_X(x) \, dx.$$
Expectation

**Definition:** The **expectation** of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx.$$  

**Justification:** Say $X = n\delta$ w.p. $f_X(n\delta)\delta$ for $n \in \mathbb{Z}$. Then,

$$E[X] = \sum_n (n\delta) \Pr[X = n\delta] = \sum_n (n\delta)f_X(n\delta)\delta = \int_{-\infty}^{\infty} xf_X(x)\,dx.$$  

Indeed, for any $g$, one has $\int g(x)\,dx \approx \sum_n g(n\delta)\delta$. 
Expectation

Definition: The expectation of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx.$$ 

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$ for $n \in \mathbb{Z}$. Then,

$$E[X] = \sum_{n} (n\delta) \Pr[X = n\delta] = \sum_{n} (n\delta) f_X(n\delta)\delta = \int_{-\infty}^{\infty} x f_X(x) dx.$$ 

Indeed, for any $g$, one has $\int g(x) dx \approx \sum_{n} g(n\delta)\delta$. Choose $g(x) = xf_X(x)$. 

Definition: The expectation of a random variable $X$ with pdf $f(x)$ is defined as

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) \, dx.$$ 

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$ for $n \in \mathbb{Z}$. Then,

$$E[X] = \sum_n (n\delta) Pr[X = n\delta] = \sum_n (n\delta)f_X(n\delta)\delta = \int_{-\infty}^{\infty} xf_X(x) \, dx.$$ 

Indeed, for any $g$, one has $\int g(x) \, dx \approx \sum_n g(n\delta)\delta$. Choose $g(x) = xf_X(x)$. 

![Graph of expectation and expectation function](image)
Examples of Expectation

1. $X = U[0,1]$. Then, $f_{X}(x) = 1_{0 \leq x \leq 1}$. Thus, $E[X] = \int_{-\infty}^{\infty} x f_{X}(x) \, dx = \int_{0}^{1} x \cdot 1 \, dx = \frac{1}{2}$.

2. $X =$ distance to 0 of dart shot uniformly in unit circle. Then $f_{X}(x) = 2x 1_{0 \leq x \leq 1}$. Thus, $E[X] = \int_{-\infty}^{\infty} x f_{X}(x) \, dx = \int_{0}^{1} x \cdot 2x \, dx = \frac{2}{3}$.
Examples of Expectation

1. $X = U[0,1]$. 

     Then, 
     $f_X(x) = \begin{cases} 0 & \text{if } 0 \leq x \leq 1 \\ 1 & \text{otherwise} \end{cases}$.

     Thus, 
     $E[X] = \int_{\infty}^{\infty} x f_X(x) \, dx = \int_{0}^{1} x \cdot 1 \, dx = \left[ \frac{x^2}{2} \right]_{0}^{1} = \frac{1}{2}$.

2. $X =$ distance to 0 of dart shot uniformly in unit circle.

     Then 
     $f_X(x) = 2x \cdot 1 \{0 \leq x \leq 1\}$.

     Thus, 
     $E[X] = \int_{\infty}^{\infty} x f_X(x) \, dx = \int_{0}^{1} x \cdot 2x \, dx = \left[ \frac{2x^3}{3} \right]_{0}^{1} = \frac{2}{3}$. 
Examples of Expectation

1. $X = U[0, 1]$. Then, $f_X(x) =$
Examples of Expectation

1. \( X = U[0, 1] \). Then, \( f_X(x) = 1\{0 \leq x \leq 1\} \).
Examples of Expectation

1. $X = U[0,1]$. Then, $f_X(x) = 1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) \, dx$$
Examples of Expectation

1. $X = U[0, 1]$. Then, $f_X(x) = 1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx = \int_{0}^{1} x \cdot 1 dx =$$
1. \( X = U[0,1] \). Then, \( f_X(x) = 1\{0 \leq x \leq 1\} \). Thus,

\[
E[X] = \int_{-\infty}^{\infty} xf_X(x) \, dx = \int_{0}^{1} x \cdot 1 \, dx = \left[ \frac{x^2}{2} \right]_{0}^{1} = \]


Examples of Expectation

1. $X = U[0, 1]$. Then, $f_X(x) = 1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx = \int_{0}^{1} x \cdot 1\,dx = \left[ \frac{x^2}{2} \right]_0^1 = \frac{1}{2}. $$
Examples of Expectation

1. \( X = U[0, 1] \). Then, \( f_X(x) = 1\{0 \leq x \leq 1\} \). Thus,

\[
E[X] = \int_{-\infty}^{\infty} x f_X(x) \, dx = \int_{0}^{1} x \cdot 1 \, dx = \left[ \frac{x^2}{2} \right]_{0}^{1} = \frac{1}{2}.
\]

2. \( X = \) distance to 0 of dart shot uniformly in unit circle.
Examples of Expectation

1. $X = U[0,1]$. Then, $f_X(x) = 1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x) \, dx = \int_{0}^{1} x \cdot 1 \, dx = \left[ \frac{x^2}{2} \right]_0^1 = \frac{1}{2}.$$ 

2. $X =$ distance to 0 of dart shot uniformly in unit circle. Then $f_X(x) = 2x1\{0 \leq x \leq 1\}.$
Examples of Expectation

1. $X = U[0, 1]$. Then, $f_X(x) = 1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx = \int_{0}^{1} x \cdot 1\,dx = \left[ \frac{x^2}{2} \right]_{0}^{1} = \frac{1}{2}.$$

2. $X =$ distance to 0 of dart shot uniformly in unit circle. Then $f_X(x) = 2x1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx$$
Examples of Expectation

1. $X = U[0, 1]$. Then, $f_X(x) = 1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx = \int_{0}^{1} x \cdot 1 dx = \left[ \frac{x^2}{2} \right]_{0}^{1} = \frac{1}{2}.$$

2. $X =$ distance to 0 of dart shot uniformly in unit circle. Then $f_X(x) = 2x \cdot 1\{0 \leq x \leq 1\}$. Thus,

$$E[X] = \int_{-\infty}^{\infty} xf_X(x)dx = \int_{0}^{1} x \cdot 2x dx =$$
Examples of Expectation

1. \( X = U[0, 1] \). Then, \( f_X(x) = 1\{0 \leq x \leq 1\} \). Thus,

\[
E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx = \int_{0}^{1} x \cdot 1\,dx = \left[ \frac{x^2}{2} \right]_{0}^{1} = \frac{1}{2}.
\]

2. \( X = \) distance to 0 of dart shot uniformly in unit circle. Then \( f_X(x) = 2x1\{0 \leq x \leq 1\} \). Thus,

\[
E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx = \int_{0}^{1} x \cdot 2x\,dx = \left[ \frac{2x^3}{3} \right]_{0}^{1} = \frac{2}{3}.
\]
Examples of Expectation

1. $X = U[0, 1]$. Then, $f_X(x) = 1\{0 \leq x \leq 1\}$. Thus,

$$
E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx = \int_{0}^{1} x \cdot 1\,dx = \left[ \frac{x^2}{2} \right]_0^1 = \frac{1}{2}.
$$

2. $X =$ distance to 0 of dart shot uniformly in unit circle. Then $f_X(x) = 2x1\{0 \leq x \leq 1\}$. Thus,

$$
E[X] = \int_{-\infty}^{\infty} xf_X(x)\,dx = \int_{0}^{1} x \cdot 2x\,dx = \left[ \frac{2x^3}{3} \right]_0^1 = \frac{2}{3}.
$$
Examples of Expectation

3. $X = \text{Exponential}(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$.

Thus, $E[X] = \int_0^\infty x \lambda e^{-\lambda x} dx = \left[ -\frac{xe^{-\lambda x}}{\lambda} - \frac{e^{-\lambda x}}{\lambda^2} \right]_0^\infty = 0 - 0 + \frac{1}{\lambda} = \frac{1}{\lambda}$.

Recall the integration by parts formula:

$$\int_a^b u(x) dv(x) = \left[ u(x)v(x) \right]_a^b - \int_a^b v(x) du(x).$$

Thus, $\int_0^\infty xe^{-\lambda x} dx = \left[ -\frac{xe^{-\lambda x}}{\lambda} - \frac{e^{-\lambda x}}{\lambda^2} \right]_0^\infty = 0 - 0 + \frac{1}{\lambda} = \frac{1}{\lambda}$.
Examples of Expectation

3. $X = Expo(\lambda)$. 
Examples of Expectation

3. $X = Expo(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$. 
Examples of Expectation

3. $X = Expo(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x}1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x \lambda e^{-\lambda x} \, dx$$
Examples of Expectation

3. $X = \text{Expo}(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x}1\{x \geq 0\}$. Thus,

$$E[X] = \int_{0}^{\infty} x \lambda e^{-\lambda x} dx = -\int_{0}^{\infty} x de^{-\lambda x}.$$
Examples of Expectation

3. $X = Expo(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x}1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x\lambda e^{-\lambda x}\,dx = -\int_0^\infty xde^{-\lambda x}.$$

Recall the integration by parts formula:

$$\int_a^b u(x)dv(x) = [u(x)v(x)]_a^b - \int_a^b v(x)du(x)$$
Examples of Expectation

3. $X = \text{Expo}(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x \lambda e^{-\lambda x} \, dx = - \int_0^\infty xde^{-\lambda x}.$$ 

Recall the integration by parts formula:

$$\int_a^b u(x)dv(x) = [u(x)v(x)]_a^b - \int_a^b v(x)du(x)$$

$$= u(b)v(b) - u(a)v(a) - \int_a^b v(x)du(x).$$
Examples of Expectation

3. $X = \text{Expo}(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x}1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x\lambda e^{-\lambda x} \, dx = -\int_0^\infty xde^{-\lambda x}.$$ 

Recall the integration by parts formula:

$$\int_a^b u(x)dv(x) = [u(x)v(x)]_a^b - \int_a^b v(x)du(x)$$

$$= u(b)v(b) - u(a)v(a) - \int_a^b v(x)du(x).$$

Thus,

$$\int_0^\infty xde^{-\lambda x} = [xe^{-\lambda x}]_0^\infty - \int_0^\infty e^{-\lambda x} \, dx$$
Examples of Expectation

3. \( X = \text{Expo}(\lambda) \). Then, \( f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\} \). Thus,

\[
E[X] = \int_{0}^{\infty} x \lambda e^{-\lambda x} \, dx = -\int_{0}^{\infty} xde^{-\lambda x}.
\]

Recall the integration by parts formula:

\[
\int_{a}^{b} u(x)dv(x) = [u(x)v(x)]_{a}^{b} - \int_{a}^{b} v(x)du(x)
\]

\[
= u(b)v(b) - u(a)v(a) - \int_{a}^{b} v(x)du(x).
\]

Thus,

\[
\int_{0}^{\infty} xde^{-\lambda x} = [xe^{-\lambda x}]_{0}^{\infty} - \int_{0}^{\infty} e^{-\lambda x} \, dx
\]

\[
= 0 - 0 + \frac{1}{\lambda} \int_{0}^{\infty} de^{-\lambda x} = \frac{1}{\lambda}.
\]
Examples of Expectation

3. $X = \text{Expo}(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x \lambda e^{-\lambda x} \, dx = -\int_0^\infty x e^{-\lambda x} \, dx.$$ 

Recall the integration by parts formula:

$$\int_a^b u(x) \, dv(x) = \left[ u(x)v(x) \right]_a^b - \int_a^b v(x) \, du(x)$$

Thus,

$$\int_0^\infty x \, de^{-\lambda x} = \left[ xe^{-\lambda x} \right]_0^\infty - \int_0^\infty e^{-\lambda x} \, dx$$

$$= 0 - 0 + \frac{1}{\lambda} \int_0^\infty de^{-\lambda x} = \frac{1}{\lambda}. $$
Examples of Expectation

3. $X = \text{Expo}(\lambda)$. Then, $f_X(x) = \lambda e^{-\lambda x}1\{x \geq 0\}$. Thus,

$$E[X] = \int_0^\infty x\lambda e^{-\lambda x} \, dx = -\int_0^\infty xde^{-\lambda x}.$$ 

Recall the integration by parts formula:

$$\int_a^b u(x)dv(x) = \left[ u(x)v(x) \right]_a^b - \int_a^b v(x)du(x)$$

Thus,

$$\int_0^\infty xde^{-\lambda x} = \left[ xe^{-\lambda x} \right]_0^\infty - \int_0^\infty e^{-\lambda x} \, dx$$

$$= 0 - 0 + \frac{1}{\lambda} \int_0^\infty de^{-\lambda x} = -\frac{1}{\lambda}.$$ 

Hence, $E[X] = \frac{1}{\lambda}$. 
Independent Continuous Random Variables

Definition: The continuous RVs $X$ and $Y$ are independent if
\[\Pr[X \in A, Y \in B] = \Pr[X \in A] \Pr[Y \in B], \forall A, B.\]

Theorem: The continuous RVs $X$ and $Y$ are independent if and only if
\[f_{X,Y}(x,y) = f_X(x)f_Y(y).\]

Proof: As in the discrete case.

Definition: The continuous RVs $X_1, \ldots, X_n$ are mutually independent if
\[\Pr[X_1 \in A_1, \ldots, X_n \in A_n] = \Pr[X_1 \in A_1] \cdots \Pr[X_n \in A_n], \forall A_1, \ldots, A_n.\]

Theorem: The continuous RVs $X_1, \ldots, X_n$ are mutually independent if and only if
\[f_{X_1,\ldots,X_n}(x_1,\ldots,x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).\]

Proof: As in the discrete case.
Independent Continuous Random Variables

Definition:

Theorem:

Proof:

Definition:

Theorem:

Proof:
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$\Pr[ X \in A, Y \in B ] = \Pr[ X \in A ] \Pr[ Y \in B ], \quad \forall A, B.$$

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x, y) = f_X(x) f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$\Pr[ X_1 \in A_1, \ldots, X_n \in A_n ] = \Pr[ X_1 \in A_1 ] \cdots \Pr[ X_n \in A_n ], \quad \forall A_1, \ldots, A_n.$$

**Theorem:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if and only if

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

**Proof:** As in the discrete case.
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:**
**Independent Continuous Random Variables**

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A] Pr[Y \in B], \forall A, B.$$  

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

**Proof:**
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:**
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$  

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$Pr[X_1 \in A_1, \ldots, X_n \in A_n] = Pr[X_1 \in A_1]\cdots Pr[X_n \in A_n], \forall A_1, \ldots, A_n.$$
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$ 

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$Pr[X_1 \in A_1, \ldots, X_n \in A_n] = Pr[X_1 \in A_1] \cdots Pr[X_n \in A_n], \forall A_1, \ldots, A_n.$$ 

**Theorem:**
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x, y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$Pr[X_1 \in A_1, \ldots, X_n \in A_n] = Pr[X_1 \in A_1] \cdots Pr[X_n \in A_n], \forall A_1, \ldots, A_n.$$ 

**Theorem:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if and only if
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$Pr[X_1 \in A_1, \ldots, X_n \in A_n] = Pr[X_1 \in A_1] \cdots Pr[X_n \in A_n], \forall A_1, \ldots, A_n.$$ 

**Theorem:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if and only if

$$f_X(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A]Pr[Y \in B], \forall A, B.$$ 

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$Pr[X_1 \in A_1, \ldots, X_n \in A_n] = Pr[X_1 \in A_1] \ldots Pr[X_n \in A_n], \forall A_1, \ldots, A_n.$$ 

**Theorem:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if and only if

$$f_X(x_1, \ldots, x_n) = f_{X_1}(x_1) \ldots f_{X_n}(x_n).$$

**Proof:**
Independent Continuous Random Variables

**Definition:** The continuous RVs $X$ and $Y$ are independent if

$$Pr[X \in A, Y \in B] = Pr[X \in A] Pr[Y \in B], \forall A, B.$$  

**Theorem:** The continuous RVs $X$ and $Y$ are independent if and only if

$$f_{X,Y}(x,y) = f_X(x)f_Y(y).$$

**Proof:** As in the discrete case.

**Definition:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if

$$Pr[X_1 \in A_1, \ldots, X_n \in A_n] = Pr[X_1 \in A_1] \cdots Pr[X_n \in A_n], \forall A_1, \ldots, A_n.$$  

**Theorem:** The continuous RVs $X_1, \ldots, X_n$ are mutually independent if and only if

$$f_X(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

**Proof:** As in the discrete case.
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm. They agree they will wait for 10 minutes. What is the probability they meet? Here, \((X, Y)\) are the times when the friends reach the restaurant. The shaded area are the pairs where \(|X - Y| < \frac{1}{6}\), i.e., such that they meet. The complement is the sum of two rectangles. When you put them together, they form a square with sides \(\frac{5}{6}\). Thus, \(\Pr[\text{meet}] = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}\).
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes.
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant.

The shaded area are the pairs where \(|X - Y| < \frac{1}{6}\), i.e., such that they meet.
The complement is the sum of two rectangles. When you put them together, they form a square with sides \(\frac{5}{6}\).

Thus, \(\Pr[\text{meet}] = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}\).
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant.

The shaded area are the pairs where \(|X - Y| < 1/6\),
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant.

The shaded area are the pairs where \(|X - Y| < 1/6\), i.e., such that they meet.
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant.

The shaded area are the pairs where \(|X - Y| < 1/6\), i.e., such that they meet.

The complement is the sum of two rectangles.

\[
\Pr[\text{meet}] = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}.
\]
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant.

The shaded area are the pairs where \(|X - Y| < 1/6\), i.e., such that they meet.

The complement is the sum of two rectangles. When you put them together, they form a square.
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.
They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant.

The shaded area are the pairs where \(|X - Y| < 1/6\), i.e., such that they meet.

The complement is the sum of two rectangles. When you put them together, they form a square with sides 5/6.
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm. They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant. The shaded area are the pairs where \(|X - Y| < 1/6\), i.e., such that they meet.

The complement is the sum of two rectangles. When you put them together, they form a square with sides \(5/6\).

Thus, \(Pr[\text{meet}] = 1 - \left(\frac{5}{6}\right)^2 = \ldots\)
Meeting at a Restaurant

Two friends go to a restaurant independently uniformly at random between noon and 1pm.

They agree they will wait for 10 minutes. What is the probability they meet?

Here, \((X, Y)\) are the times when the friends reach the restaurant.

The shaded area are the pairs where \(|X - Y| < 1/6\), i.e., such that they meet.

The complement is the sum of two rectangles. When you put them together, they form a square with sides 5/6.

Thus, \(Pr[\text{meet}] = 1 - \left(\frac{5}{6}\right)^2 = \frac{11}{36}\).
Breaking a Stick

You break a stick at two points chosen independently uniformly at random. What is the probability you can make a triangle with the three pieces?

Let $X, Y$ be the two break points along the $[0, 1]$ stick. A triangle if $A < B + C$, $B < A + C$, and $C < A + B$.

If $X < Y$, this means $X < 0.5$, $Y < X + 0.5$, $Y > 0.5$. This is the blue triangle.

If $X > Y$, get red triangle, by symmetry. Thus, $\Pr[\text{make triangle}] = \frac{1}{4}$. 
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X$, $Y$ be the two break points along the $[0, 1]$ stick.
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X$, $Y$ be the two break points along the $[0,1]$ stick.

A triangle if

Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X, Y$ be the two break points along the $[0, 1]$ stick.

A triangle if
$A < B + C, B < A + C, \text{ and } C < A + B.$

If $X < Y$, this means
$X < 0.5,$
You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X, Y$ be the two break points along the $[0, 1]$ stick.

A triangle if

$A < B + C, B < A + C, \text{ and } C < A + B.$

If $X < Y$, this means

$X < 0.5, Y < X + .5,$
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X, Y$ be the two break points along the $[0, 1]$ stick.

A triangle if


If $X < Y$, this means

$X < 0.5, Y < X + .5, Y > 0.5$. 
You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let \( X, Y \) be the two break points along the \([0, 1]\) stick.

A triangle if
\[
A < B + C, \quad B < A + C, \quad \text{and} \quad C < A + B.
\]

If \( X < Y \), this means
\[
X < 0.5, \quad Y < X + 0.5, \quad Y > 0.5.
\]
This is the blue triangle.
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X, Y$ be the two break points along the $[0, 1]$ stick.

A triangle if

$A < B + C, B < A + C, \text{ and } C < A + B.$

If $X < Y$, this means

$X < 0.5, \ Y < X + .5, \ Y > 0.5.$

This is the blue triangle.

If $X > Y$, get red triangle, by symmetry.
You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X, Y$ be the two break points along the $[0, 1]$ stick.

A triangle if $A < B + C, B < A + C, \text{ and } C < A + B$.

If $X < Y$, this means $X < 0.5, Y < X + .5, Y > 0.5$.

This is the blue triangle.

If $X > Y$, get red triangle, by symmetry.
Breaking a Stick

You break a stick at two points chosen independently uniformly at random.

What is the probability you can make a triangle with the three pieces?

Let $X, Y$ be the two break points along the $[0, 1]$ stick.

A triangle if $A < B + C, B < A + C,$ and $C < A + B.$

If $X < Y,$ this means $X < 0.5, Y < X + .5, Y > 0.5.$

This is the blue triangle.

If $X > Y,$ get red triangle, by symmetry.

Thus, $Pr[\text{make triangle}] = 1/4.$
Maximum of Two Exponentials

Let $X = \text{Exp}(\lambda)$ and $Y = \text{Exp}(\mu)$ be independent. Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has $\Pr[Z < z] = \Pr[X < z, Y < z] = \Pr[X < z] \Pr[Y < z] = (1 - e^{-\lambda z})(1 - e^{-\mu z}) = 1 - e^{-\lambda z} - e^{-\mu z} + e^{-(\lambda + \mu)z}$.

Thus, $f_Z(z) = \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu) e^{-(\lambda + \mu)z}$, $\forall z > 0$.

Since, $\int_0^\infty x \lambda e^{-\lambda x} dx = \lambda \left[-xe^{-\lambda x}\right]_0^\infty + \lambda e^{-\lambda x}|_0^\infty = \frac{\lambda}{\lambda^2} = 1/\lambda$.

$E[Z] = \int_0^\infty z f_Z(z) dz = 1/\lambda + 1/\mu - 1/\lambda + 1/\mu.$
Maximum of Two Exponentials

Let $X = Expo(\lambda)$ and $Y = Expo(\mu)$ be independent.
Maximum of Two Exponentials

Let $X = Expo(\lambda)$ and $Y = Expo(\mu)$ be independent.
Define $Z = \max\{X, Y\}$.
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent. Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$. 
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent.
Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.
Maximum of Two Exponentials

Let $X = Expo(\lambda)$ and $Y = Expo(\mu)$ be independent. Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent. Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

$$Pr[Z < z] = Pr[X < z, Y < z]$$
Maximum of Two Exponentials

Let \( X = \text{Expo}(\lambda) \) and \( Y = \text{Expo}(\mu) \) be independent.
Define \( Z = \max\{X, Y\} \).

Calculate \( E[Z] \).
We compute \( f_Z \), then integrate.

One has

\[
Pr[Z < z] = Pr[X < z, Y < z] = Pr[X < z]Pr[Y < z]
\]
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent. Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

$$Pr[Z < z] = Pr[X < z, Y < z] = Pr[X < z]Pr[Y < z]$$

$$= (1 - e^{-\lambda z})(1 - e^{-\mu z}) =$$
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent.
Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

$$Pr[Z < z] = Pr[X < z, Y < z] = Pr[X < z]Pr[Y < z]$$
$$= (1 - e^{-\lambda z})(1 - e^{-\mu z}) = 1 - e^{-\lambda z} - e^{-\mu z} + e^{-(\lambda+\mu)z}$$
Maximum of Two Exponentials

Let $X = \text{Exp}(\lambda)$ and $Y = \text{Exp}(\mu)$ be independent.
Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

$$Pr[Z < z] = Pr[X < z, Y < z] = Pr[X < z]Pr[Y < z]$$
$$= (1 - e^{-\lambda z})(1 - e^{-\mu z}) = 1 - e^{-\lambda z} - e^{-\mu z} + e^{-(\lambda + \mu)z}$$

Thus,

$$f_Z(z) = \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu) e^{-(\lambda + \mu)z}, \forall z > 0.$$
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent.

Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

$$
Pr[Z < z] = Pr[X < z, Y < z] = Pr[X < z]Pr[Y < z]
= (1 - e^{-\lambda z})(1 - e^{-\mu z}) = 1 - e^{-\lambda z} - e^{-\mu z} + e^{-(\lambda + \mu)z}
$$

Thus,

$$
f_Z(z) = \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu)e^{-(\lambda + \mu)z}, \forall z > 0.
$$

Since,

$$
\int_0^\infty x \lambda e^{-\lambda x} dx = \lambda \left[ -\frac{xe^{-\lambda x}}{\lambda} - \frac{e^{-\lambda x}}{\lambda^2} \right]_0^\infty = \frac{1}{\lambda}.
$$

$$
E[Z] = \int_0^\infty zf_Z(z)dz =
$$
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent.

Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

$$Pr[Z < z] = Pr[X < z, Y < z] = Pr[X < z] Pr[Y < z]$$

$$= (1 - e^{-\lambda z})(1 - e^{-\mu z}) = 1 - e^{-\lambda z} - e^{-\mu z} + e^{-(\lambda + \mu) z}$$

Thus,

$$f_Z(z) = \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu) e^{-(\lambda + \mu) z}, \forall z > 0.$$  

Since, $\int_0^\infty x \lambda e^{-\lambda x} dx = \lambda [\frac{-xe^{-\lambda x}}{\lambda} - \frac{e^{-\lambda x}}{\lambda^2}]_0^\infty = \frac{1}{\lambda}$.

$$E[Z] = \int_0^\infty zf_Z(z)dz = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}.$$
Maximum of Two Exponentials

Let $X = \text{Expo}(\lambda)$ and $Y = \text{Expo}(\mu)$ be independent.
Define $Z = \max\{X, Y\}$.

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

$$Pr[Z < z] = Pr[X < z, Y < z] = Pr[X < z]Pr[Y < z] = (1 - e^{-\lambda z})(1 - e^{-\mu z}) = 1 - e^{-\lambda z} - e^{-\mu z} + e^{-(\lambda+\mu)z}$$

Thus,

$$f_Z(z) = \lambda e^{-\lambda z} + \mu e^{-\mu z} - (\lambda + \mu) e^{-(\lambda+\mu)z}, \forall z > 0.$$ 

Since, $\int_0^\infty x e^{-\lambda x} dx = \lambda \left[-\frac{xe^{-\lambda x}}{\lambda} - \frac{e^{-\lambda x}}{\lambda^2}\right]_0^\infty = \frac{1}{\lambda}$.

$$E[Z] = \int_0^\infty z f_Z(z) dz = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}. $$
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. Exponential(1).

Define $Z = \max \{X_1, X_2, \ldots, X_n\}$.

Calculate $E[Z]$.

We use a recursion.

The key idea is as follows:

$Z = \min \{X_1, \ldots, X_n\} + V$

where $V$ is the maximum of $n - 1$ i.i.d. Exponential(1).

This follows from the memoryless property of the exponential.

Let then $A_n = E[Z]$.

We see that $A_n = E[\min \{X_1, \ldots, X_n\}] + A_{n-1} = \frac{1}{n} + A_{n-1}$ because the minimum of Exponential is Exponential with the sum of the rates.

Hence, $E[Z] = A_n = \frac{1}{n} + \frac{1}{2} + \cdots + \frac{1}{n} = H(n)$. 
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $Expo(1)$. 
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $Expo(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$. 

Calculate $E[Z]$. We use a recursion. The key idea is as follows:

$Z = \min\{X_1, \ldots, X_n\} + V$

where $V$ is the maximum of $n-1$ i.i.d. $Expo(1)$. This follows from the memoryless property of the exponential.

Let then $A_n = E[Z]$. We see that $A_n = E[\min\{X_1, \ldots, X_n\}] + A_{n-1} = 1/n + A_{n-1}$ because the minimum of Expo is Expo with the sum of the rates. Hence, $E[Z] = A_n = 1 + 1/2 + \cdots + 1/n = H(n)$. 


Maximum of \( n \) i.i.d. Exponentials

Let \( X_1, \ldots, X_n \) be i.i.d. Expo(1). Define \( Z = \max\{X_1, X_2, \ldots, X_n\} \). Calculate \( E[Z] \).
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $\text{Expo}(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$.

Calculate $E[Z]$.

We use a recursion.
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. Expo(1). Define $Z = \max\{X_1, X_2, \ldots, X_n\}$. Calculate $E[Z]$.

We use a recursion. The key idea is as follows:
Maximum of \( n \) i.i.d. Exponentials

Let \( X_1, \ldots, X_n \) be i.i.d. \( \text{Expo}(1) \). Define \( Z = \max\{X_1, X_2, \ldots, X_n\} \).

Calculate \( E[Z] \).

We use a recursion. The key idea is as follows:

\[
Z = \min\{X_1, \ldots, X_n\} + V
\]

where \( V \) is the maximum of \( n - 1 \) i.i.d. \( \text{Expo}(1) \).
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $Expo(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$.

Calculate $E[Z]$.

We use a recursion. The key idea is as follows:

$$Z = \min\{X_1, \ldots, X_n\} + V$$

where $V$ is the maximum of $n-1$ i.i.d. $Expo(1)$. This follows from the memoryless property of the exponential.
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $\text{Expo}(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$.

Calculate $E[Z]$.

We use a recursion. The key idea is as follows:

$$Z = \min\{X_1, \ldots, X_n\} + V$$

where $V$ is the maximum of $n - 1$ i.i.d. $\text{Expo}(1)$. This follows from the memoryless property of the exponential.

Let then $A_n = E[Z]$. 
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $\text{Expo}(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$.

Calculate $E[Z]$.

We use a recursion. The key idea is as follows:

$$Z = \min\{X_1, \ldots, X_n\} + V$$

where $V$ is the maximum of $n-1$ i.i.d. $\text{Expo}(1)$. This follows from the memoryless property of the exponential.

Let then $A_n = E[Z]$. We see that

$$A_n = E[\min\{X_1, \ldots, X_n\}] + A_{n-1}$$
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $\text{Expo}(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$. Calculate $E[Z]$.

We use a recursion. The key idea is as follows:

$$Z = \min\{X_1, \ldots, X_n\} + V$$

where $V$ is the maximum of $n - 1$ i.i.d. $\text{Expo}(1)$. This follows from the memoryless property of the exponential.

Let then $A_n = E[Z]$. We see that

$$A_n = E[\min\{X_1, \ldots, X_n\}] + A_{n-1} = \frac{1}{n} + A_{n-1}$$
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $\text{Expo}(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$. Calculate $E[Z]$.

We use a recursion. The key idea is as follows:

$$Z = \min\{X_1, \ldots, X_n\} + V$$

where $V$ is the maximum of $n - 1$ i.i.d. $\text{Expo}(1)$. This follows from the memoryless property of the exponential.

Let then $A_n = E[Z]$. We see that

$$A_n = E[\min\{X_1, \ldots, X_n\}] + A_{n-1} = \frac{1}{n} + A_{n-1}$$

because the minimum of $\text{Expo}$ is $\text{Expo}$ with the sum of the rates.
Maximum of $n$ i.i.d. Exponentials

Let $X_1, \ldots, X_n$ be i.i.d. $\text{Expo}(1)$. Define $Z = \max\{X_1, X_2, \ldots, X_n\}$. Calculate $E[Z]$.

We use a recursion. The key idea is as follows:

$$Z = \min\{X_1, \ldots, X_n\} + V$$

where $V$ is the maximum of $n-1$ i.i.d. $\text{Expo}(1)$. This follows from the memoryless property of the exponential.

Let then $A_n = E[Z]$. We see that

$$A_n = E[\min\{X_1, \ldots, X_n\}] + A_{n-1} = \frac{1}{n} + A_{n-1}$$

because the minimum of $\text{Expo}$ is $\text{Expo}$ with the sum of the rates. Hence,

$$E[Z] = A_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = H(n).$$
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits. This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

Model:

\[ X = U[0, 1] \]

is the continuous value. \( Y \) is the closest multiple of \( 2^{-n} \) to \( X \). Thus, we can represent \( Y \) with \( n \) bits. The error is \( Z = X - Y \). The power of the noise is \( E[Z^2] \).

Analysis:

We see that \( Z \) is uniform in \([0, a = 2^{-n} - 1] \). Thus,

\[ E[Z^2] = a^2 \int_0^a dz = 1 \]

The power of the signal \( X \) is \( E[X^2] = 1 \).
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.
This introduces an error
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.
This introduces an error perceived as noise: the quantization noise.
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

Model:
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

Model: \( X = U[0,1] \) is the continuous value.
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** \( X = U[0, 1] \) is the continuous value. \( Y \) is the closest multiple of \( 2^{-n} \) to \( X \).
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits. This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** $X = U[0, 1]$ is the continuous value. $Y$ is the closest multiple of $2^{-n}$ to $X$. Thus, we can represent $Y$ with $n$ bits.
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** \( X = U[0, 1] \) is the continuous value. \( Y \) is the closest multiple of \( 2^{-n} \) to \( X \). Thus, we can represent \( Y \) with \( n \) bits. The error is \( Z := X - Y \).
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** $X = U[0, 1]$ is the continuous value. $Y$ is the closest multiple of $2^{-n}$ to $X$. Thus, we can represent $Y$ with $n$ bits. The error is $Z := X - Y$.

The power of the noise is $E[Z^2]$. 
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

Model: \( X = U[0, 1] \) is the continuous value. \( Y \) is the closest multiple of \( 2^{-n} \) to \( X \). Thus, we can represent \( Y \) with \( n \) bits. The error is \( Z := X - Y \).

The power of the noise is \( E[Z^2] \).

Analysis:
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits. This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** \( X = U[0, 1] \) is the continuous value. \( Y \) is the closest multiple of \( 2^{-n} \) to \( X \). Thus, we can represent \( Y \) with \( n \) bits. The error is \( Z := X - Y \).

The power of the noise is \( E[Z^2] \).

**Analysis:** We see that \( Z \) is uniform in \([0, a = 2^{-(n+1)}]\).
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** $X = U[0, 1]$ is the continuous value. $Y$ is the closest multiple of $2^{-n}$ to $X$. Thus, we can represent $Y$ with $n$ bits. The error is $Z := X - Y$.

The power of the noise is $E[Z^2]$.

**Analysis:** We see that $Z$ is uniform in $[0, a = 2^{-(n+1)}]$. Thus,

$$E[Z^2] = \frac{a^2}{3} = \frac{1}{3} 2^{-2(n+1)}.$$
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** $X = U[0, 1]$ is the continuous value. $Y$ is the closest multiple of $2^{-n}$ to $X$. Thus, we can represent $Y$ with $n$ bits. The error is $Z := X - Y$.

The power of the noise is $E[Z^2]$.

**Analysis:** We see that $Z$ is uniform in $[0, a = 2^{-(n+1)}]$.

Thus,

$$E[Z^2] = \frac{a^2}{3} = \frac{1}{3} 2^{-2(n+1)}.$$

The power of the signal $X$ is $E[X^2] =$
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

**Model:** $X = U[0, 1]$ is the continuous value. $Y$ is the closest multiple of $2^{-n}$ to $X$. Thus, we can represent $Y$ with $n$ bits. The error is $Z := X - Y$.

The power of the noise is $E[Z^2]$.

**Analysis:** We see that $Z$ is uniform in $[0, a = 2^{-(n+1)}]$.

Thus,

$$E[Z^2] = \frac{a^2}{3} = \frac{1}{3} 2^{-2(n+1)}.$$

The power of the signal $X$ is $E[X^2] = \frac{1}{3}$.
Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

Model: $X = U[0, 1]$ is the continuous value. $Y$ is the closest multiple of $2^{-n}$ to $X$. Thus, we can represent $Y$ with $n$ bits. The error is $Z := X - Y$.

The power of the noise is $E[Z^2]$.

Analysis: We see that $Z$ is uniform in $[0, a = 2^{-(n+1)}]$.

Thus,

$$E[Z^2] = \frac{a^2}{3} = \frac{1}{3} 2^{-2(n+1)}.$$

The power of the signal $X$ is $E[X^2] = \frac{1}{3}$. 
Quantization Noise

We saw that

\[ E[Z^2] = \frac{1}{3} - 2(n + 1) \]

and

\[ E[X^2] = \frac{1}{3}. \]

The signal to noise ratio (SNR) is the power of the signal divided by the power of the noise. Thus,

\[ \text{SNR} = \frac{2}{2(n + 1)}. \]

Expressed in decibels, one has

\[ \text{SNR (dB)} = 10\log_{10} \left( \frac{\text{SNR}}{\frac{1}{2}} \right) = 20(n + 1) \log_{10}(2) \approx 6(n + 1). \]

For instance, if \( n = 16 \), then \( \text{SNR (dB)} \approx 112 \text{ dB} \).
Quantization Noise

We saw that $E[Z^2] = \frac{1}{3} 2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$. 
We saw that $E[Z^2] = \frac{1}{3} 2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

The **signal to noise ratio** (SNR)
We saw that $E[Z^2] = \frac{1}{3}2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

The **signal to noise ratio** (SNR) is the power of the signal divided by the power of the noise.
We saw that $E[Z^2] = \frac{1}{3} 2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

The **signal to noise ratio** (SNR) is the power of the signal divided by the power of the noise.

Thus, 

$$SNR = 2^{2(n+1)}.$$
Quantization Noise

We saw that $E[Z^2] = \frac{1}{3} 2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

The **signal to noise ratio** (SNR) is the power of the signal divided by the power of the noise.

Thus,

$$SNR = 2^{2(n+1)}.$$ 

Expressed in decibels, one has

$$SNR(dB) = 10 \log_{10}(SNR).$$
We saw that $E[Z^2] = \frac{1}{3}2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

The **signal to noise ratio** (SNR) is the power of the signal divided by the power of the noise.

Thus,

$$SNR = 2^{2(n+1)}.$$

Expressed in decibels, one has

$$SNR(dB) = 10\log_{10}(SNR) = 20(n+1)\log_{10}(2)$$
We saw that $E[Z^2] = \frac{1}{3}2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

The **signal to noise ratio** (SNR) is the power of the signal divided by the power of the noise.

Thus,

$$SNR = 2^{2(n+1)}.$$

Expressed in decibels, one has

$$SNR(dB) = 10\log_{10}(SNR) = 20(n+1)\log_{10}(2) \approx 6(n+1).$$

For instance, if $n = 16$, then $SNR(dB) \approx 112$ dB.
Quantization Noise

We saw that $E[Z^2] = \frac{1}{3}2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

The **signal to noise ratio** (SNR) is the power of the signal divided by the power of the noise.

Thus,

$$SNR = 2^{2(n+1)}.$$  

Expressed in decibels, one has

$$SNR(dB) = 10 \log_{10}(SNR) = 20(n+1) \log_{10}(2) \approx 6(n+1).$$

For instance, if $n = 16$, then $SNR(dB) \approx 112 dB$. 

Problem 1: Pick two points \(X\) and \(Y\) independently and uniformly at random in \([0, 1]\). What is \(E[(X - Y)^2]\)?

Analysis: One has \(E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]\) = \(\frac{1}{3} + \frac{1}{3} - \frac{2}{12} = \frac{1}{6}\).

Problem 2: What about in a unit square?

Analysis: One has \(E[||X - Y||^2] = E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2]\) = \(2 \times \frac{1}{6}\).

Problem 3: What about in \(n\) dimensions?

\(n\)-dimensional.
Expected Squared Distance

Problem 1:

Pick two points $X$ and $Y$ independently and uniformly at random in $\left[0, 1\right]$. What is $E[(X - Y)^2]$?

Analysis:

One has $E[(X - Y)^2] = E[X^2] + E[Y^2] - 2E[XY] = \frac{1}{3} + \frac{1}{3} - \frac{1}{2} = \frac{1}{6}$.

Problem 2:

What about in a unit square?

Analysis:

One has $E[|X - Y|^2] = E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2] = 2E[(X - Y)^2] = \frac{2}{3}$.

Problem 3:

What about in $n$ dimensions?

$n \frac{6}{n}$.
Expected Squared Distance

Problem 1: Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$. 
**Problem 1:** Pick two points \( X \) and \( Y \) independently and uniformly at random in \([0, 1]\).

What is \( E[(X - Y)^2] \)?
Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0,1]$.

What is $E[(X - Y)^2]$?

**Analysis:**

...
Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

What is $E[(X - Y)^2]$?

**Analysis:** One has

$$E[(X - Y)^2] =$$
Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

What is $E[(X - Y)^2]$?

**Analysis:** One has

\[
E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]
\]
Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$. What is $E[(X - Y)^2]$?

**Analysis:** One has

$$E[(X - Y)^2] = E[X^2 + Y^2 - 2XY] = \frac{1}{3} + \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2}.$$
Expected Squared Distance

**Problem 1:** Pick two points \( X \) and \( Y \) independently and uniformly at random in \([0, 1]\).

What is \( E[(X - Y)^2] \)?

**Analysis:** One has

\[
E[(X - Y)^2] = E[X^2 + Y^2 - 2XY] \\
= \frac{1}{3} + \frac{1}{3} - 2 \frac{1}{2} \frac{1}{2} \\
= 2 \frac{1}{3} - \frac{1}{2} = \frac{1}{6}.
\]
Expected Squared Distance

Problem 1: Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

What is $E[(X - Y)^2]$?

Analysis: One has

$$E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]$$
$$= \frac{1}{3} + \frac{1}{3} - 2\frac{1}{2} \frac{1}{2}$$
$$= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.$$ 

Problem 2:
Expected Squared Distance

Problem 1: Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.
What is $E[(X - Y)^2]$?

Analysis: One has

$$E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]$$
$$= \frac{1}{3} + \frac{1}{3} - 2\frac{1}{2}\frac{1}{2}$$
$$= 2\frac{1}{3} - \frac{1}{2} = \frac{1}{6}.$$ 

Problem 2: What about in a unit square?
Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.
What is $E[(X - Y)^2]$?

**Analysis:** One has

$$ E[(X - Y)^2] = E[X^2 + Y^2 - 2XY] $$
$$ = \frac{1}{3} + \frac{1}{3} - 2\frac{1}{2}\frac{1}{2} $$
$$ = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. $$

**Problem 2:** What about in a unit square?

**Analysis:**

$$ E[(X - Y)^2] = 2 \times \frac{1}{6} = \frac{1}{3}. $$
**Expected Squared Distance**

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

What is $E[(X - Y)^2]$?

**Analysis:** One has

\[
E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]
\]

\[
= \frac{1}{3} + \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2}
\]

\[
= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.
\]

**Problem 2:** What about in a unit square?

**Analysis:** One has

\[
E[||X - Y||^2] =
\]
Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$. What is $E[(X - Y)^2]$?

**Analysis:** One has

$$E[(X - Y)^2] = E[X^2 + Y^2 - 2XY] = \frac{1}{3} + \frac{1}{3} - 2 \frac{1}{2} \frac{1}{2} = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.$$

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$E[\|X - Y\|^2] = E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2]$$
Expected Squared Distance

Problem 1: Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

What is $E[(X - Y)^2]$?

Analysis: One has

$$E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]$$

$$= \frac{1}{3} + \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2}$$

$$= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.$$

Problem 2: What about in a unit square?

Analysis: One has

$$E[||X - Y||^2] = E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2]$$

$$= 2 \times \frac{1}{6}.$$
**Expected Squared Distance**

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

What is $E[(X - Y)^2]$?

**Analysis:** One has

\[
E[(X - Y)^2] = E[X^2 + Y^2 - 2XY] = \frac{1}{3} + \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.
\]

**Problem 2:** What about in a unit square?

**Analysis:** One has

\[
E[\|X - Y\|^2] = E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2] = 2 \times \frac{1}{6}.
\]

**Problem 3:**
Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

What is $E[(X - Y)^2]$?

**Analysis:** One has

$$E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]$$

$$= \frac{1}{3} + \frac{1}{3} - 2 \times \frac{1}{2 \times 2}$$

$$= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.$$ 

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$E[||X - Y||^2] = E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2]$$

$$= 2 \times \frac{1}{6}.$$ 

**Problem 3:** What about in $n$ dimensions?
Expected Squared Distance

**Problem 1:** Pick two points \( X \) and \( Y \) independently and uniformly at random in \([0, 1]\).

What is \( E[(X - Y)^2]\)?

**Analysis:** One has

\[
E[(X - Y)^2] = E[X^2 + Y^2 - 2XY]
\]

\[
= \frac{1}{3} + \frac{1}{3} - 2 \cdot \frac{1}{2} \cdot \frac{1}{2}
\]

\[
= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}.
\]

**Problem 2:** What about in a unit square?

**Analysis:** One has

\[
E[\|X - Y\|^2] = E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2]
\]

\[
= 2 \times \frac{1}{6}.
\]

**Problem 3:** What about in \( n \) dimensions? \( \frac{n}{6} \).
Geometric and Exponential

The geometric and exponential distributions are similar. They are both memoryless. Consider flipping a coin every $\frac{1}{N}$ second with $\Pr[H] = \frac{p}{N}$, where $N \gg 1$. Let $X$ be the time until the first $H$. Fact: $X \approx \text{Expo}(p)$. Analysis: Note that $\Pr[X > t] \approx \Pr[\text{first } Nt \text{ flips are tails}] = (1 - \frac{p}{N})^N \approx \exp\{-pt\}$. Indeed, $(1 - aN)^N \approx \exp\{-a\}$. 
The geometric and exponential distributions are similar.
The geometric and exponential distributions are similar. They are both memoryless.
Geometric and Exponential

The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$. Let $X$ be the time until the first $H$.

Fact: $X \approx \text{Expo}(p)$.

Analysis: Note that $Pr[X > t] \approx Pr[\text{first } Nt \text{ flips are tails}] = (1 - p/N)^{Nt} \approx \exp\{-pt\}$.

Indeed, $(1 - aN)^N \approx \exp\{-a\}$. 
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$. 

Let $X$ be the time until the first $H$.

Fact: $X \approx \text{Expo}(p)$.

Analysis: Note that $Pr[X > t] \approx Pr[\text{first } Nt \text{ flips are tails}] = (1 - p/N)^{Nt} \approx \exp\{ -pt \}$.

Indeed, $(1 - a/N)^N \approx \exp\{ -a \}$. 

Geometric and Exponential
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$. 
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$.

**Fact:**

$$Pr[X > t] \approx Pr[\text{first } Nt \text{ flips are tails}] = \left(1 - \frac{p}{N}\right)^{Nt} \approx \exp\{-pt\}.$$
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$.

**Fact:** $X \approx Expo(p)$. 

Geometric and Exponential

The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$.

**Fact:** $X \approx Expo(p)$.

**Analysis:**
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$.

**Fact:** $X \approx Expo(p)$.

**Analysis:** Note that

$$Pr[X > t] \approx Pr[\text{first } Nt \text{ flips are tails}]$$
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$.

**Fact:** $X \approx Expo(p)$.

**Analysis:** Note that

$$
Pr[X > t] \approx Pr[\text{first } Nt \text{ flips are tails}]
\approx (1 - \frac{p}{N})^{Nt}
$$
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$.

Fact: $X \approx \text{Expo}(p)$.

Analysis: Note that

$$Pr[X > t] \approx Pr[\text{first } Nt \text{ flips are tails}] = (1 - \frac{p}{N})^{Nt} \approx \exp\{-pt\}.$$
The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$, where $N \gg 1$.

Let $X$ be the time until the first $H$.

**Fact:** $X \approx \text{Expo}(p)$.

**Analysis:** Note that

$$Pr[X > t] \approx Pr[\text{first } Nt \text{ flips are tails}]$$

$$= (1 - \frac{p}{N})^{Nt} \approx \exp\{-pt\}.$$

Indeed, $(1 - \frac{a}{N})^N \approx \exp\{-a\}$. 
Summary

Continuous Probability

▶ Continuous RVs are essentially the same as discrete RVs
▶ Think that $X \approx x$ with probability $f_X(x) \in \mathbb{R}$
▶ Sums become integrals, ....
▶ The exponential distribution is magical: memoryless.
Continuous Probability

Continuous RVs are essentially the same as discrete RVs.

Think that

$$X \approx x$$

with probability

$$f_X(x) \in$$

Sums become integrals, ...

The exponential distribution is magical: memoryless.
Continuous Probability

- Continuous RVs are essentially the same as discrete RVs

The exponential distribution is magical: memoryless.
Continuous Probability

- Continuous RVs are essentially the same as discrete RVs
- Think that $X \approx x$ with probability $f_X(x)\epsilon$
Continuous Probability

- Continuous RVs are essentially the same as discrete RVs
- Think that $X \approx x$ with probability $f_X(x)\epsilon$
- Sums become integrals, ....
Continuous Probability

- Continuous RVs are essentially the same as discrete RVs
- Think that $X \approx x$ with probability $f_X(x) \varepsilon$
- Sums become integrals, ....
- The exponential distribution is magical:
Continuous Probability

- Continuous RVs are essentially the same as discrete RVs
- Think that $X \approx x$ with probability $f_X(x) \epsilon$
- Sums become integrals, ...
- The exponential distribution is magical: memoryless.