# Probability

# Probability

Probability!

# Probability

Probability!
Confuses us.

# Probability

Probability!
Confuses us. But really neat.

# Probability

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

# Probability

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

# Probability

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

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.

Sample Space:$\Omega$, $Pr[\omega]$.

# Probability

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

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

# 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 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: $f(x) = \frac{dF(x)}{dx}$.
$\int_{-\infty}^{\infty} f(x) = 1$.

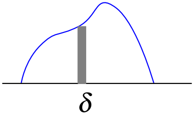
# Probability

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: $Pr[X = x]$
$\sum_x Pr[X = x] = 1$.

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

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



# Probability Rules are all good.

Conditional Probabity.

# Probability Rules are all good.

Conditional Probabity.

Events: $A, B$

# Probability Rules are all good.

Conditional Probabity.

Events: $A, B$

Discrete: “Heads”, “Tails”, $X = 1$, $Y = 5$.

# Probability Rules are all good.

Conditional Probabity.

Events: $A, B$

Discrete: "Heads", "Tails", $X = 1$, $Y = 5$.

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

# Probability Rules are all good.

Conditional Probabity.

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

# Probability Rules are all good.

Conditional Probabity.

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

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

# Probability Rules are all good.

Conditional Probabity.

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

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

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

# Probability Rules are all good.

Conditional Probabity.

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

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

# Probability Rules are all good.

Conditional Probabity.

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

$Pr$[“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.

# Probability Rules are all good.

Conditional Probabity.

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

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

# Probability Rules are all good.

Conditional Probabity.

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

$Pr$["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 $X \in [0, .5]$

# Probability Rules are all good.

Conditional Probabity.

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: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$
$Pr[\text{"Second Heads"}] = Pr[HH] + Pr[HT]$
$B$ 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 Probabity.

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

$Pr[$"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 $X \in [0, .5]$

Product Rule: $Pr[A \cap B] = Pr[A|B]Pr[B]$.
Bayes Rule: $Pr[A|B] = Pr[B|A]Pr[A]/Pr[B]$.

# Probability Rules are all good.

Conditional Probabity.

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

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

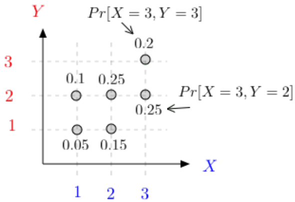
Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$
$Pr[$“Second Heads”$] = Pr[HH] + Pr[HT]$
$B$ is First coin heads.
$Pr[X \in [.45, .55]] = Pr[X \in [.45, .50]] + Pr[X \in (.50, .55]]$
$B$ is $X \in [0, .5]$

Product Rule: $Pr[A \cap B] = Pr[A|B] Pr[B]$.
Bayes Rule: $Pr[A|B] = Pr[B|A] Pr[A] / Pr[B]$.

All work for continuous with intervals as events.

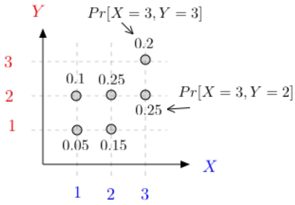
# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |



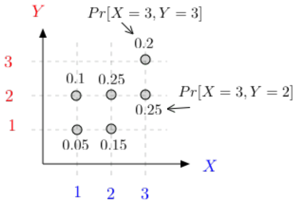
# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |



# Joint distribution.

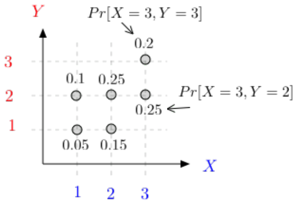
| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




Marginal Distribution?

# Joint distribution.

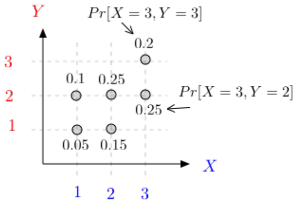
| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




Marginal Distribution? Here is one.

# Joint distribution.

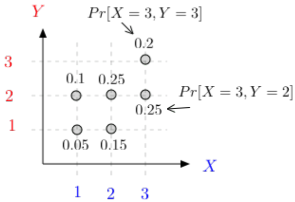
| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

# Joint distribution.

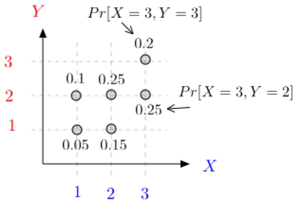
| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |

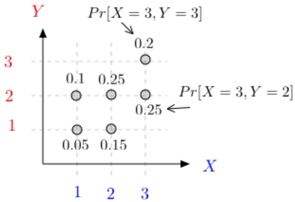



Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

$E[Y|X]$?

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |



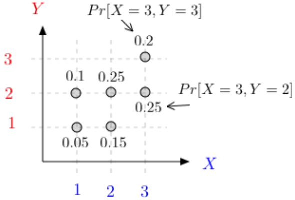

Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

$E[Y|X]$?

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

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |



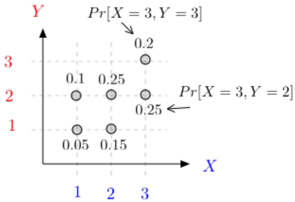

Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

$E[Y|X]$?

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

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |



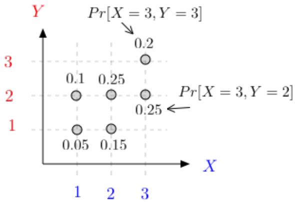

Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

$E[Y|X]$?

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

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |



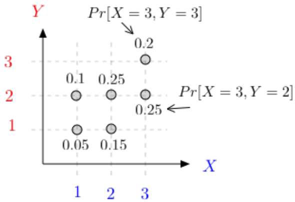

Marginal Distribution? Here is one. And here is another.
The distribution of one of the variables.

$E[Y|X]$?

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

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

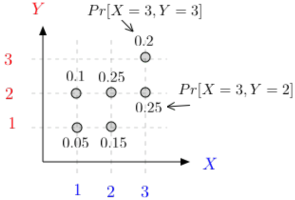
$E[Y|X]$?

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

$E[Y]$

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

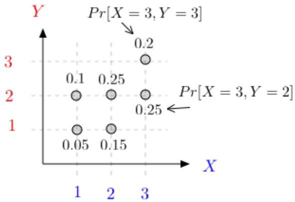
$E[Y|X]$?

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

$E[Y] = E[E[Y|X]] =$

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

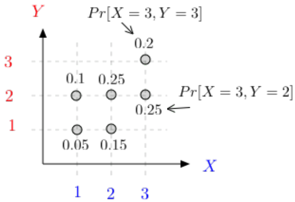
$E[Y|X]$?

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

$E[Y] = E[E[Y|X]] = E[Y|X=1]Pr[X=1]$

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

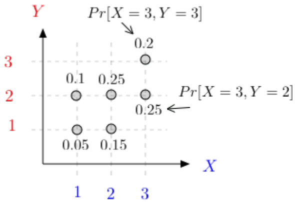
$E[Y|X]$?

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

$E[Y] = E[E[Y|X]] = E[Y|X=1]Pr[X=1] + E[Y|X=2]Pr[X=2]$

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

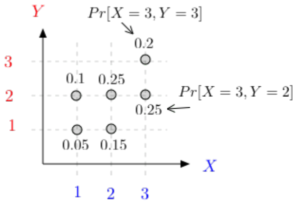
$E[Y|X]$?

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

$$E[Y] = E[E[Y|X]] = E[Y|X=1]Pr[X=1] + E[Y|X=2]Pr[X=2] + \cdots$$

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

$E[Y|X]$?

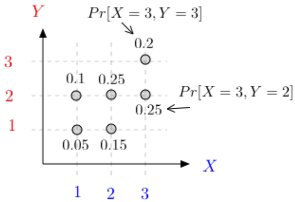
$$E[Y|X=1] = (.03\times 1 + .2\times 2 + .21\times 3 + .02\times 5)/.44 = \frac{1.16}{.44}.$$
$$E[Y|X=2] = (.05\times 1 + .01\times 2 + .06\times 3 + .2\times 5)/.32 = \frac{1.25}{.32}.$$
$$E[Y|X=4] = (.1\times 1 + .03\times 2 + .03\times 3 + .02\times 5)/.18 = \frac{.35}{.18}.$$
$$E[Y|X=8] = (.02\times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

$$E[Y] = E[E[Y|X]] = E[Y|X=1]Pr[X=1] + E[Y|X=2]Pr[X=2] + \cdots$$

$$E[Y] = (1.16 + 1.25 + .35 + .10) = 2.86.$$

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

$E[Y|X]$?

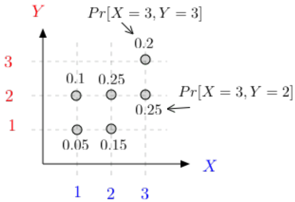
$$E[Y|X=1] = (.03\times 1 + .2\times 2 + .21\times 3 + .02\times 5)/.44 = \frac{1.16}{.44}.$$
$$E[Y|X=2] = (.05\times 1 + .01\times 2 + .06\times 3 + .2\times 5)/.32 = \frac{1.25}{.32}.$$
$$E[Y|X=4] = (.1\times 1 + .03\times 2 + .03\times 3 + .02\times 5)/.18 = \frac{.35}{.18}.$$
$$E[Y|X=8] = (.02\times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

$$E[Y] = E[E[Y|X]] = E[Y|X=1]Pr[X=1] + E[Y|X=2]Pr[X=2] + \cdots$$

$$E[Y] = (1.16 + 1.25 + .35 + .10) = 2.86.$$

# Joint distribution.

| Y/X | 1 | 2 | 4 | 8 | |
|---|---|---|---|---|---|
| 1 | .03 | .05 | .1 | .02 | .20 |
| 2 | .2 | .01 | .03 | .02 | .26 |
| 3 | .21 | .06 | .03 | .01 | .31 |
| 5 | .02 | .2 | .02 | .01 | .25 |
| | .44 | .32 | .18 | .06 | |




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

$E[Y|X]$?

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

$$E[Y] = E[E[Y|X]] = E[Y|X=1]Pr[X=1] + E[Y|X=2]Pr[X=2] + \cdots$$

$$E[Y] = (1.16 + 1.25 + .35 + .10) = 2.86.$$

# Multiple Continuous Random Variables

# 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 $x, y \in \Re$

# Multiple Continuous Random Variables

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

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

# Multiple Continuous Random Variables

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

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

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 $x,y \in \Re$ where

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

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 $x, y \in \Re$ where

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

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 $x,y \in \Re$ where

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

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 $A \subset \Re^2$.

# Multiple Continuous Random Variables

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

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

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 $A \subset \Re^2$. Then

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

# Multiple Continuous Random Variables

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

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

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 $A \subset \Re^2$. 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 $x, y \in \Re$ where

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

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 $A \subset \Re^2$. 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 $x,y \in \Re$ where

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

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 $A \subset \Re^2$. 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 $x,y \in \Re$ where

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

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 $A \subset \Re^2$. 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 $x,y \in \Re$ where

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

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 $A \subset \Re^2$. 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 $x, y \in \Re$ where

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

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 $A \subset \Re^2$. 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:** $\mathbf{X} = (X_1, \ldots, X_n)$ with $f_{\mathbf{X}}(\mathbf{x})$.

# Example of Continuous $(X, Y)$

# Example of Continuous $(X, Y)$

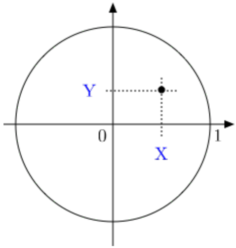
Pick a point $(X, Y)$ uniformly in the unit circle.

# Example of Continuous $(X, Y)$

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

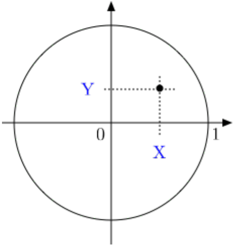
# Example of Continuous $(X, Y)$

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



# Example of Continuous $(X, Y)$

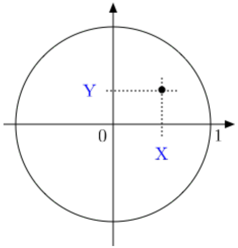
Pick a point $(X, Y)$ uniformly in the unit circle.




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

# Example of Continuous $(X, Y)$

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



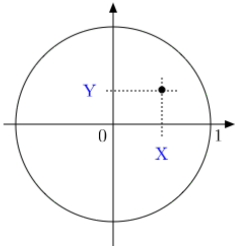

$$\implies f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.$$

Some events!

$$Pr[X > 0, Y > 0] =$$

# Example of Continuous $(X, Y)$

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



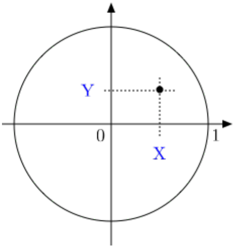

$$\implies f_{X,Y}(x,y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}.$$

Some events!

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

# Example of Continuous $(X, Y)$

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




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

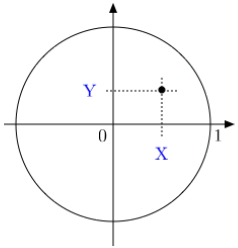
Some events!

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

$$Pr[X < 0, Y > 0] =$$

# Example of Continuous $(X, Y)$

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




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

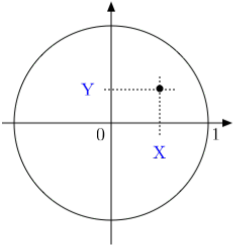
Some events!

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

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

# 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!

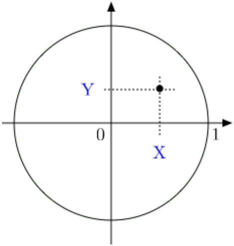
$$Pr[X > 0, Y > 0] = \frac{1}{4}$$

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

$$Pr[X^2 + Y^2 \leq r^2] =$$

# 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!

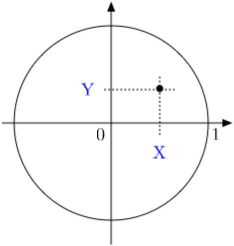
$$Pr[X > 0, Y > 0] = \frac{1}{4}$$

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

$$Pr[X^2 + Y^2 \leq r^2] = r^2$$

# Example of Continuous $(X, Y)$

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




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

Some events!

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

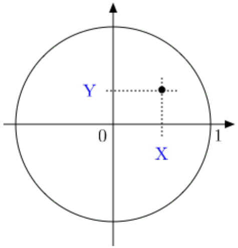
$$Pr[X < 0, Y > 0] = \frac{1}{4}$$

$$Pr[X^2 + Y^2 \leq r^2] = r^2$$

$$Pr[X > Y] =$$

# 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!

$$Pr[X > 0, Y > 0] = \frac{1}{4}$$
$$Pr[X < 0, Y > 0] = \frac{1}{4}$$
$$Pr[X^2 + Y^2 \leq r^2] = r^2$$
$$Pr[X > Y] = \frac{1}{2}.$$

# Example of Continuous $(X, Y)$

# Example of Continuous $(X, Y)$

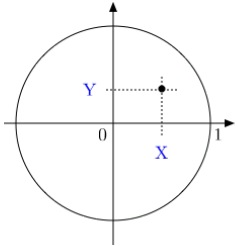
Pick a point $(X, Y)$ uniformly in the unit circle.

# Example of Continuous $(X, Y)$

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

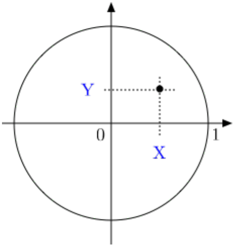
# Example of Continuous $(X, Y)$

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



# Example of Continuous $(X, Y)$

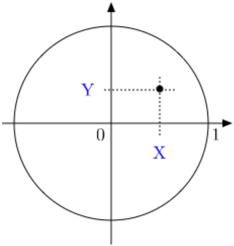
Pick a point $(X, Y)$ uniformly in the unit circle.




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

# Example of Continuous $(X, Y)$

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

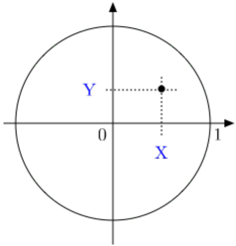



$f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}$.

Marginals?

# Example of Continuous $(X, Y)$

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



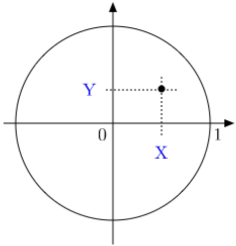

$f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}$.

Marginals?

$f_X(x) =$

# Example of Continuous $(X, Y)$

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



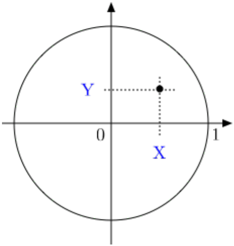

$f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}$.

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.



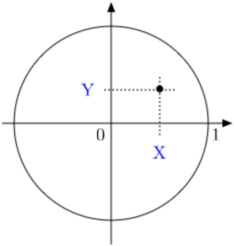

$f_{X,Y}(x, y) = \frac{1}{\pi} 1\{x^2 + y^2 \leq 1\}$.

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.




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

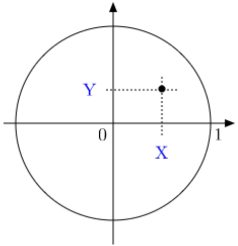
Marginals?

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

$f_Y(y) =$

# Example of Continuous $(X, Y)$

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




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

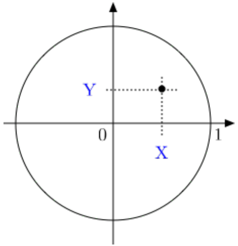
Marginals?

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

$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx$

# 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?

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

$f_Y(y) = \int_{-\infty}^{\infty} f_{X,Y}(x, y) dx = \frac{2}{\pi}\sqrt{1 - y^2}$

# *Expo*($\lambda$)

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

# $Expo(\lambda)$

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

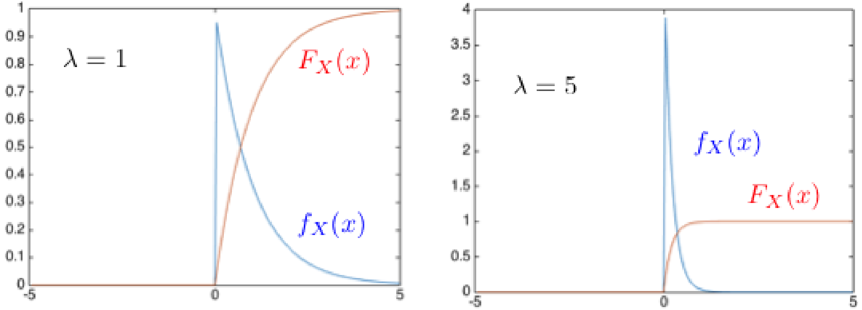
$$f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$$

# *Expo*($\lambda$)

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

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

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

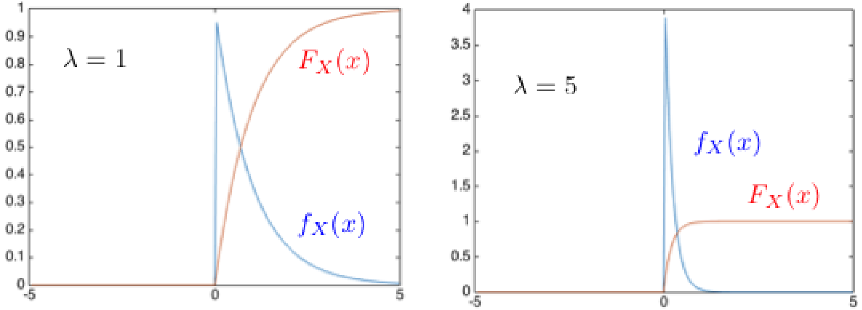


# $Expo(\lambda)$

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

# 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 $s, t > 0$,

# Some Properties

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

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

# 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 $s, t > 0$,

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

# Some Properties

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

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

# Some Properties

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

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

# Some Properties

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

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

'Used is a good as new.'

# Some Properties

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

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

'Used is a good as new.'

**2. Scaling** ***Expo*****.**

# Some Properties

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

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

'Used is a good as new.'

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

# Some Properties

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

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

'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 $s, t > 0$,

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

'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 $s, t > 0$,

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

'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 $X = Expo(\lambda)$. Then, for $s, t > 0$,

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

'Used is a good as new.'

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

$$
\begin{aligned}
Pr[Y > t] &= Pr[aX > t] = Pr[X > t/a] \\
&= e^{-\lambda(t/a)} = e^{-(\lambda/a)t} =
\end{aligned}
$$

# Some Properties

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

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

'Used is a good as new.'

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

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

# Some Properties

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

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

'Used is a good as new.'

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

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

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

# Some Properties

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

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

'Used is a good as new.'

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

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

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

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

# More Properties

# More Properties

**3. Scaling Uniform.**

# More Properties

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

# More Properties

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

$$Pr[Y \in (y, y+\delta)] \quad = \quad Pr[a + bX \in (y, y+\delta)] =$$

# More Properties

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

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

# More Properties

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

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

# More Properties

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

$$\begin{aligned} Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = \frac{1}{b}\delta, \text{ for} \end{aligned}$$

# More Properties

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

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

# More Properties

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

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

# More Properties

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

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

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

# More Properties

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

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

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

# More Properties

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

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

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

# Some More Properties

**4. Scaling pdf.**

# Some More Properties

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

# Some More Properties

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

$$Pr[Y \in (y, y+\delta)] \quad = \quad Pr[a + bX \in (y, y+\delta)] =$$

# Some More Properties

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

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

# Some More Properties

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

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

# Some More Properties

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

$$\begin{aligned} Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = f_X(\frac{y-a}{b})\frac{\delta}{b}. \end{aligned}$$

# Some More Properties

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

$$
\begin{aligned}
Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\
&= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = f_X(\frac{y-a}{b})\frac{\delta}{b}.
\end{aligned}
$$

Now, the left-hand side is

# Some More Properties

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

$$\begin{aligned} Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = f_X(\frac{y-a}{b})\frac{\delta}{b}. \end{aligned}$$

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

# Some More Properties

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

$$\begin{aligned} Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = f_X(\frac{y-a}{b})\frac{\delta}{b}. \end{aligned}$$

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

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

# Some More Properties

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

$$\begin{aligned} Pr[Y \in (y, y+\delta)] &= Pr[a + bX \in (y, y+\delta)] = Pr[X \in (\frac{y-a}{b}, \frac{y+\delta-a}{b})] \\ &= Pr[X \in (\frac{y-a}{b}, \frac{y-a}{b} + \frac{\delta}{b})] = f_X(\frac{y-a}{b})\frac{\delta}{b}. \end{aligned}$$

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

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

# Expectation

**Definition:**

# Expectation

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

# 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

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

**Justification:**

# Expectation

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

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

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

# 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,

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

# 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,

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

# 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,

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

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

# 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 $g(x) = x f_X(x)$.

# Expectation

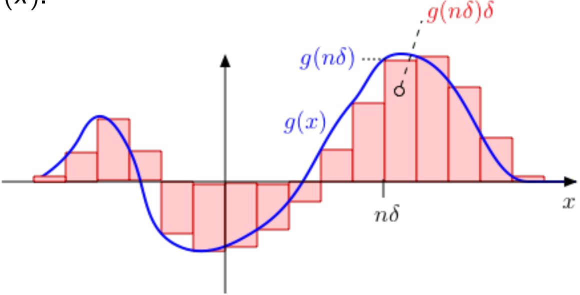
**Definition:** The **expectation** of a random variable $X$ with pdf $f(x)$ is *defined* as

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

**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 $g(x) = x f_X(x)$.



# Examples of Expectation

# Examples of Expectation

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

# 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,

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

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx =$$

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 =$$

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

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

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

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

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

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

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

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \cdot 2x dx =$$

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \cdot 2x dx = [\frac{2x^3}{3}]_0^1 =$$

# Examples of Expectation

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x.1 dx = [\frac{x^2}{2}]_0^1 = \frac{1}{2}.$$

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

$$E[X] = \int_{-\infty}^{\infty} x f_X(x) dx = \int_0^1 x \cdot 2x dx = [\frac{2x^3}{3}]_0^1 = \frac{2}{3}.$$

# Examples of Expectation

# 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. $X = Expo(\lambda)$. 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,

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

Recall the integration by parts formula:

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

# Examples of Expectation

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

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

Recall the integration by parts formula:

$$\begin{aligned}\int_a^b u(x) dv(x) &= \left[u(x)v(x)\right]_a^b - \int_a^b v(x) du(x) \\ &= u(b)v(b) - u(a)v(a) - \int_a^b v(x) du(x).\end{aligned}$$

# Examples of Expectation

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

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

Recall the integration by parts formula:

$$\begin{aligned} \int_a^b u(x) dv(x) &= \left[u(x)v(x)\right]_a^b - \int_a^b v(x) du(x) \\ &= u(b)v(b) - u(a)v(a) - \int_a^b v(x) du(x). \end{aligned}$$

Thus,

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

# Examples of Expectation

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

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

Recall the integration by parts formula:

$$\begin{aligned} \int_a^b u(x) dv(x) &= \left[u(x)v(x)\right]_a^b - \int_a^b v(x) du(x) \\ &= u(b)v(b) - u(a)v(a) - \int_a^b v(x) du(x). \end{aligned}$$

Thus,

$$\begin{aligned} \int_0^\infty x d e^{-\lambda x} &= [x e^{-\lambda x}]_0^\infty - \int_0^\infty e^{-\lambda x} dx \\ &= 0 - 0 + \frac{1}{\lambda} \int_0^\infty d e^{-\lambda x} = \end{aligned}$$

# Examples of Expectation

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

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

Recall the integration by parts formula:

$$\begin{aligned} \int_a^b u(x) dv(x) &= \left[u(x)v(x)\right]_a^b - \int_a^b v(x) du(x) \\ &= u(b)v(b) - u(a)v(a) - \int_a^b v(x) du(x). \end{aligned}$$

Thus,

$$\begin{aligned} \int_0^\infty x de^{-\lambda x} &= [xe^{-\lambda x}]_0^\infty - \int_0^\infty e^{-\lambda x} dx \\ &= 0 - 0 + \frac{1}{\lambda} \int_0^\infty de^{-\lambda x} = -\frac{1}{\lambda}. \end{aligned}$$

# Examples of Expectation

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

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

Recall the integration by parts formula:

$$\begin{aligned}\int_a^b u(x) dv(x) &= \left[u(x)v(x)\right]_a^b - \int_a^b v(x) du(x) \\ &= u(b)v(b) - u(a)v(a) - \int_a^b v(x) du(x).\end{aligned}$$

Thus,

$$\begin{aligned}\int_0^\infty x de^{-\lambda x} &= [xe^{-\lambda x}]_0^\infty - \int_0^\infty e^{-\lambda x} dx \\ &= 0 - 0 + \frac{1}{\lambda} \int_0^\infty de^{-\lambda x} = -\frac{1}{\lambda}.\end{aligned}$$

Hence, $E[X] = \frac{1}{\lambda}$.

# Independent Continuous Random Variables

# Independent Continuous Random Variables

**Definition:**

# Independent Continuous Random Variables

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

# Independent Continuous Random Variables

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

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

# 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

# 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

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

# 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

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

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

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

**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

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

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

# Meeting at a Restaurant

# 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

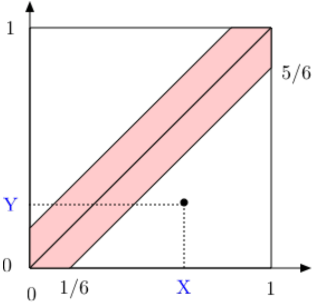
Two friends go to a restaurant independently uniformly at random between noon and 1pm.

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.

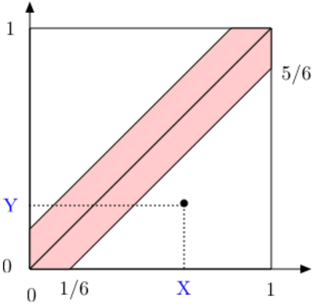
They agree they will wait for 10 minutes. What is the probability they meet?



# 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?

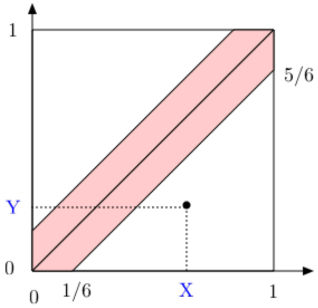



Here, $(X, Y)$ are the times when the friends reach the restaurant.

# 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?



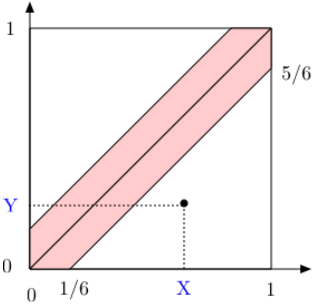

Here, $(X, Y)$ are the times when the friends reach the restaurant.

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?



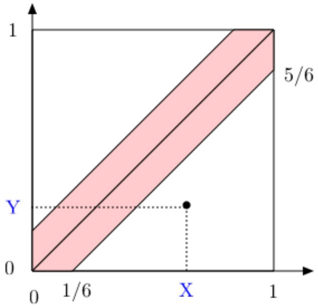

Here, $(X, Y)$ are the times when the friends reach the restaurant.

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.

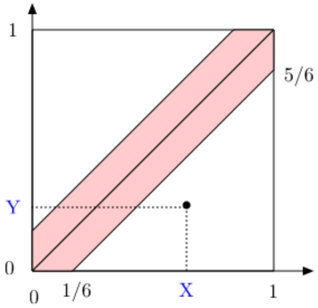
The shaded area are the pairs where $|X - Y| < 1/6$, i.e., such that they meet.

The complement is the sum of two rectangles.

# 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.

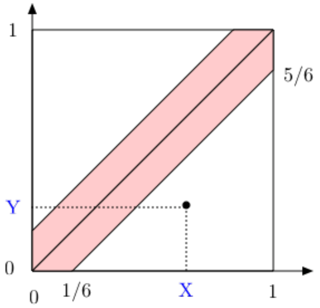
The shaded area are the pairs where $|X - Y| < 1/6$, i.e., such that they meet.

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.

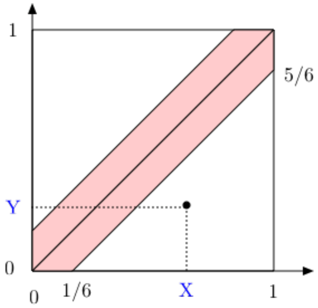
The shaded area are the pairs where $|X - Y| < 1/6$, i.e., such that they meet.

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.

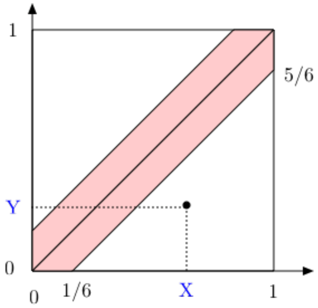
The complement is the sum of two rectangles. When you put them together, they form a square with sides $5/6$.

Thus, $Pr[\text{meet}] = 1 - (\frac{5}{6})^2 =$

# 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, $Pr[\text{meet}] = 1 - (\frac{5}{6})^2 = \frac{11}{36}$.

# Breaking a Stick

# Breaking a Stick

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

# Breaking a Stick

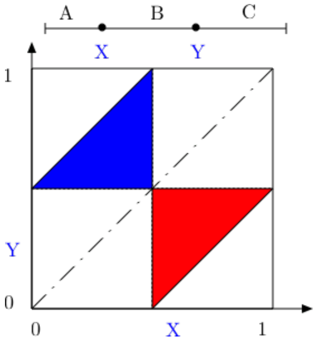
You break a stick at two points chosen independently uniformly at random.

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.

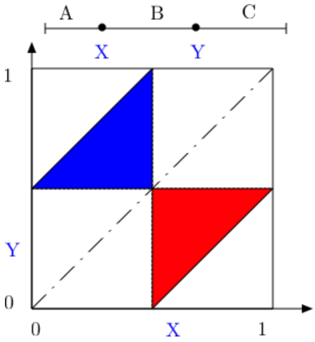
What is the probability you can make a triangle with the three pieces?



# 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?

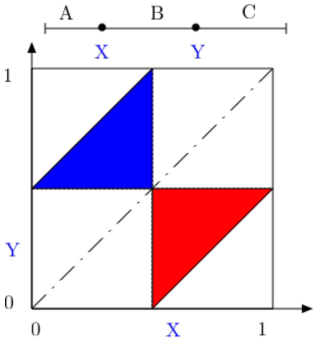



Let $X, Y$ be the two break points along the $[0, 1]$ stick.

# 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?



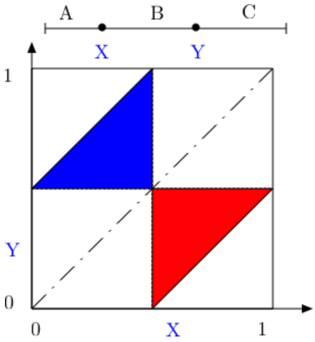

Let $X, Y$ be the two break points along the $[0,1]$ stick.

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

# 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.

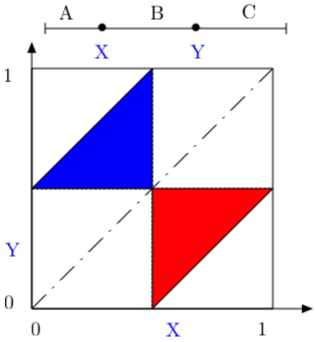
A triangle if
$A < B + C, B < A + C$, and $C < A + B$.

If $X < Y$, this means
$X < 0.5$,

# 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.

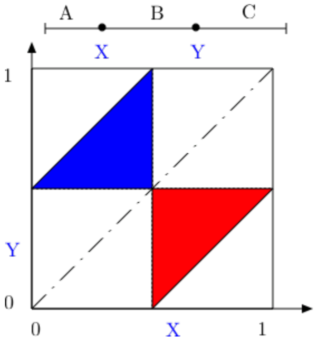
A triangle if
$A < B + C, B < A + C$, and $C < A + B$.

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

# 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.

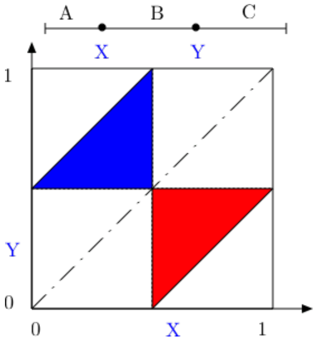
A triangle if
$A < B + C, B < A + C$, and $C < A + B$.

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

# 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.

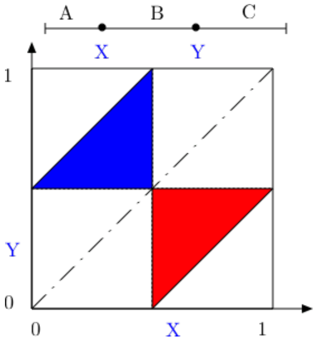
A triangle if
$A < B + C, B < A + C$, and $C < A + B$.

If $X < Y$, this means
$X < 0.5$, $Y < X + .5$, $Y > 0.5$.
This is the blue triangle.

# 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 $C < A + B$.

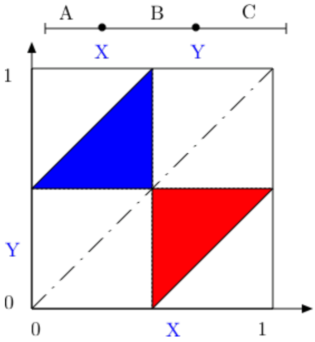
If $X < Y$, this means
$X < 0.5$, $Y < X + .5$, $Y > 0.5$.
This is the blue triangle.

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

# 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 $C < A + B$.

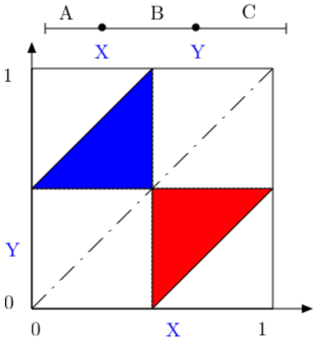
If $X < Y$, this means
$X < 0.5$, $Y < X + .5$, $Y > 0.5$.
This is the blue triangle.

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

# 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

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

Calculate $E[Z]$.

# Maximum of Two Exponentials

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

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

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

# 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

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

# 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

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

# 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

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

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

Calculate $E[Z]$.

We compute $f_Z$, then integrate.

One has

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

Thus,

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

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

$$E[Z] = \int_0^\infty z f_Z(z)dz =$$

# 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

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

Thus,

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

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

$$E[Z] = \int_0^\infty z f_Z(z) dz = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}.$$

# 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

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

Thus,

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

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

$$E[Z] = \int_0^\infty z f_Z(z) dz = \frac{1}{\lambda} + \frac{1}{\mu} - \frac{1}{\lambda + \mu}.$$

# Maximum of $n$ i.i.d. Exponentials

# 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 $Z = \max\{X_1, X_2, \ldots, X_n\}$.

# 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. *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. *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).

# 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. *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 $A_n = E[Z]$.

# 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

$$A_n = E[\min\{X_1, \ldots, X_n\}] + A_{n-1}$$

# 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

$$\begin{aligned} A_n &= E[\min\{X_1, \ldots, X_n\}] + A_{n-1} \\ &= \frac{1}{n} + A_{n-1} \end{aligned}$$

# 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

$$\begin{aligned} A_n &= E[\min\{X_1, \ldots, X_n\}] + A_{n-1} \\ &= \frac{1}{n} + A_{n-1} \end{aligned}$$

because the minimum of *Expo* is *Expo* with the sum of the rates.

# 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

$$\begin{aligned} A_n &= E[\min\{X_1, \ldots, X_n\}] + A_{n-1} \\ &= \frac{1}{n} + A_{n-1} \end{aligned}$$

because the minimum of *Expo* is *Expo* with the sum of the rates.

Hence,

$$E[Z] = A_n = 1 + \frac{1}{2} + \cdots + \frac{1}{n} = H(n).$$

# Quantization Noise

# 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

# Quantization Noise

We saw that $E[Z^2] = \frac{1}{3}2^{-2(n+1)}$ and $E[X^2] = \frac{1}{3}$.

# 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)

# 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.

# 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,

$$SNR = 2^{2(n+1)}.$$

# 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

$$SNR(dB) = 10\log_{10}(SNR)$$

# 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

$$SNR(dB) = 10\log_{10}(SNR) = 20(n+1)\log_{10}(2)$$

# 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

$$SNR(dB) = 10\log_{10}(SNR) = 20(n+1)\log_{10}(2) \approx 6(n+1).$$

# 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 $SNR(dB) \approx 112dB$.

# Expected Squared Distance

# Expected Squared Distance

**Problem 1:**

# Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0, 1]$.

# Expected Squared Distance

**Problem 1:** Pick two points $X$ and $Y$ independently and uniformly at random in $[0,1]$.

What is $E[(X-Y)^2]$?

# 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

$$E[(X - Y)^2] \quad =$$

# 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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2+Y^2-2XY] \\ &= \frac{1}{3}+\frac{1}{3}-2\frac{1}{2}\frac{1}{2} \end{aligned}$$

# 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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2+Y^2-2XY] \\ &= \frac{1}{3}+\frac{1}{3}-2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3}-\frac{1}{2}=\frac{1}{6}. \end{aligned}$$

# 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

$$
\begin{aligned}
E[(X-Y)^2] &= E[X^2+Y^2-2XY] \\
&= \frac{1}{3}+\frac{1}{3}-2\frac{1}{2}\frac{1}{2} \\
&= \frac{2}{3}-\frac{1}{2}=\frac{1}{6}.
\end{aligned}
$$

**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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2+Y^2-2XY] \\ &= \frac{1}{3}+\frac{1}{3}-2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3}-\frac{1}{2}=\frac{1}{6}. \end{aligned}$$

**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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2 + Y^2 - 2XY] \\ &= \frac{1}{3} + \frac{1}{3} - 2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

**Problem 2:** What about in a unit square?

**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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2+Y^2-2XY] \\ &= \frac{1}{3}+\frac{1}{3}-2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3}-\frac{1}{2}=\frac{1}{6}. \end{aligned}$$

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$E[||\mathbf{X}-\mathbf{Y}||^2] \quad =$$

# 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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2+Y^2-2XY] \\ &= \frac{1}{3}+\frac{1}{3}-2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3}-\frac{1}{2}=\frac{1}{6}. \end{aligned}$$

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$E[||\mathbf{X}-\mathbf{Y}||^2] = E[(X_1-Y_1)^2]+E[(X_2-Y_2)^2]$$

# 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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2 + Y^2 - 2XY] \\ &= \frac{1}{3} + \frac{1}{3} - 2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$\begin{aligned} E[||\mathbf{X} - \mathbf{Y}||^2] &= E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2] \\ &= 2 \times \frac{1}{6}. \end{aligned}$$

# 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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2+Y^2-2XY] \\ &= \frac{1}{3}+\frac{1}{3}-2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3}-\frac{1}{2}=\frac{1}{6}. \end{aligned}$$

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$\begin{aligned} E[||\mathbf{X}-\mathbf{Y}||^2] &= E[(X_1-Y_1)^2]+E[(X_2-Y_2)^2] \\ &= 2\times\frac{1}{6}. \end{aligned}$$

**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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2 + Y^2 - 2XY] \\ &= \frac{1}{3} + \frac{1}{3} - 2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$\begin{aligned} E[||\mathbf{X} - \mathbf{Y}||^2] &= E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2] \\ &= 2 \times \frac{1}{6}. \end{aligned}$$

**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

$$\begin{aligned} E[(X-Y)^2] &= E[X^2 + Y^2 - 2XY] \\ &= \frac{1}{3} + \frac{1}{3} - 2\frac{1}{2}\frac{1}{2} \\ &= \frac{2}{3} - \frac{1}{2} = \frac{1}{6}. \end{aligned}$$

**Problem 2:** What about in a unit square?

**Analysis:** One has

$$\begin{aligned} E[||\mathbf{X} - \mathbf{Y}||^2] &= E[(X_1 - Y_1)^2] + E[(X_2 - Y_2)^2] \\ &= 2 \times \frac{1}{6}. \end{aligned}$$

**Problem 3:** What about in $n$ dimensions? $\frac{n}{6}$.

# Geometric and Exponential

# Geometric and Exponential

The geometric and exponential distributions are similar.

# Geometric and Exponential

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

# Geometric and Exponential

The geometric and exponential distributions are similar. They are both memoryless.

Consider flipping a coin every $1/N$ second with $Pr[H] = p/N$,

# 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 $N \gg 1$.

# 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 $H$.

# 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:**

# 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:** $X \approx Expo(p)$.

# 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:**

# 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:** Note that

$$Pr[X > t] \quad \approx \quad Pr[\text{first } Nt \text{ flips are tails}]$$

# 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:** Note that

$$\begin{aligned} Pr[X > t] &\approx Pr[\text{first } Nt \text{ flips are tails}] \\ &= (1 - \frac{p}{N})^{Nt} \end{aligned}$$

# 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:** Note that

$$\begin{aligned} Pr[X > t] &\approx Pr[\text{first } Nt \text{ flips are tails}] \\ &= (1 - \frac{p}{N})^{Nt} \approx \exp\{-pt\}. \end{aligned}$$

# 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:** Note that

$$\begin{aligned} Pr[X > t] &\approx Pr[\text{first } Nt \text{ flips are tails}] \\ &= (1 - \frac{p}{N})^{Nt} \approx \exp\{-pt\}. \end{aligned}$$

Indeed, $(1 - \frac{a}{N})^N \approx \exp\{-a\}$.

# Summary

Continuous Probability

# Summary

Continuous Probability

# Summary

Continuous Probability

- Continuous RVs are essentially the same as discrete RVs

# Summary

**Continuous Probability**

- Continuous RVs are essentially the same as discrete RVs
- Think that $X \approx x$ with probability $f_X(x)\varepsilon$

# Summary

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, ....

# Summary

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:

# Summary

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: memoryless.