CS70: Jean Walrand: Lecture 36.

Continuous Probability 3

CS70: Jean Walrand: Lecture 36.

Continuous Probability 3

CS70: Jean Walrand: Lecture 36.

Continuous Probability 3

1. Review: CDF, PDF
2. Review: Expectation
3. Review: Independence
4. Meeting at a Restaurant
5. Breaking a Stick
6. Maximum of Exponentials
7. Quantization Noise
8. Replacing Light Bulbs
9. Expected Squared Distance
10. Geometric and Exponential

Review: CDF and PDF.

Review: CDF and PDF.

Key idea:

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$;

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0, 1]$; throw a dart in a target.

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define $Pr[\text{outcome}]$,

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define $Pr[\text{outcome}]$, then $Pr[\text{event}]$.

Instead, one starts by defining $Pr[\text{event}]$.

Thus, one defines $Pr[X \in (-\infty, x]]$

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x]$

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0, 1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon = Pr[X \in (x, x + \varepsilon)]$.

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon = Pr[X \in (x, x+\varepsilon)]$.

$F_X(\cdot)$ is the cumulative distribution function

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon = Pr[X \in (x, x+\varepsilon)]$.

$F_X(\cdot)$ is the cumulative distribution function (CDF)

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0, 1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon = Pr[X \in (x, x + \varepsilon)]$.

$F_X(\cdot)$ is the cumulative distribution function (CDF) of X.

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0, 1]$; throw a dart in a target.

Thus, one cannot define $Pr[\text{outcome}]$, then $Pr[\text{event}]$.

Instead, one starts by defining $Pr[\text{event}]$.

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon = Pr[X \in (x, x + \varepsilon)]$.

$F_X(\cdot)$ is the cumulative distribution function (CDF) of X.

$f_X(\cdot)$ is the probability density function

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0, 1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon = Pr[X \in (x, x + \varepsilon)]$.

$F_X(\cdot)$ is the cumulative distribution function (CDF) of X.

$f_X(\cdot)$ is the probability density function (PDF)

Review: CDF and PDF.

Key idea: For a continuous RV, $Pr[X = x] = 0$ for all $x \in \Re$.

Examples: Uniform in $[0,1]$; throw a dart in a target.

Thus, one cannot define Pr[outcome], then Pr[event].

Instead, one starts by defining Pr[event].

Thus, one defines $Pr[X \in (-\infty, x]] = Pr[X \leq x] =: F_X(x), x \in \Re$.

Then, one defines $f_X(x) := \frac{d}{dx} F_X(x)$.

Hence, $f_X(x)\varepsilon = Pr[X \in (x, x+\varepsilon)]$.

$F_X(\cdot)$ is the cumulative distribution function (CDF) of X.

$f_X(\cdot)$ is the probability density function (PDF) of X.

Expectation

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

Expectation

Definitions: (a) 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

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

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

(b) The expectation of a function of a random variable is defined as

Expectation

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

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

(b) The expectation of a function of a random variable is defined as

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

Expectation

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

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

(b) The expectation of a function of a random variable is defined as

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

(c) The expectation of a function of multiple random variables is defined as

Expectation

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

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

(b) The expectation of a function of a random variable is defined as

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

(c) The expectation of a function of multiple random variables is defined as

$$E[h(\mathbf{X})] = \int \cdots \int h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) dx_1 \cdots dx_n.$$

Expectation

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

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

(b) The expectation of a function of a random variable is defined as

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

(c) The expectation of a function of multiple random variables is defined as

$$E[h(\mathbf{X})] = \int \cdots \int h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) dx_1 \cdots dx_n.$$

Justifications:

Expectation

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

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

(b) The expectation of a function of a random variable is defined as

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

(c) The expectation of a function of multiple random variables is defined as

$$E[h(\mathbf{X})] = \int \cdots \int h(\mathbf{x}) f_{\mathbf{X}}(\mathbf{x}) dx_1 \cdots dx_n.$$

Justifications: Think of the discrete approximations of the continuous RVs.

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

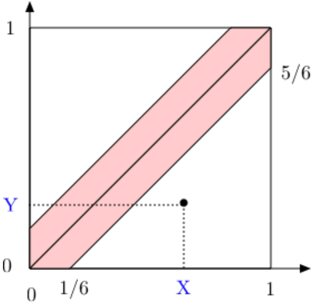
Two friends go to a restaurant independently uniformly at random between noon and 1pm.

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.

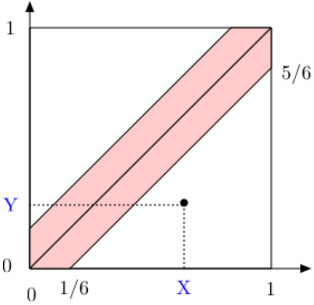
They agree they will wait for 10 minutes. What is the probability they meet?



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?

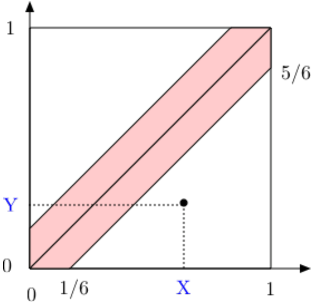


Here, (X, Y) are the times when the friends reach the restaurant.

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?



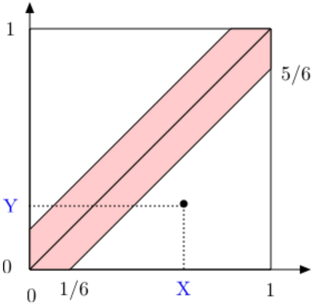
Here, (X, Y) are the times when the friends reach the restaurant.

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?



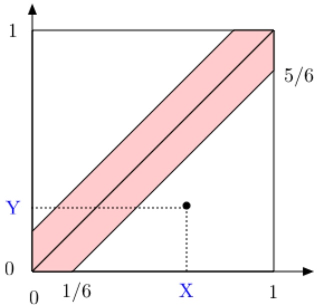
Here, (X, Y) are the times when the friends reach the restaurant.

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.

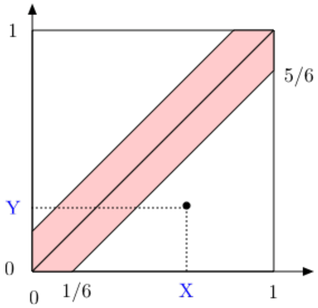
The shaded area are the pairs where $|X - Y| < 1/6$, i.e., such that they meet.

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.

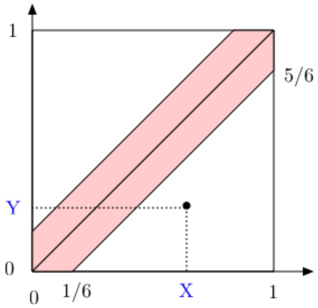
The shaded area are the pairs where $|X - Y| < 1/6$, i.e., such that they meet.

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.

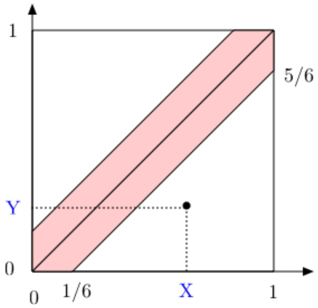
The shaded area are the pairs where $|X - Y| < 1/6$, i.e., such that they meet.

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.

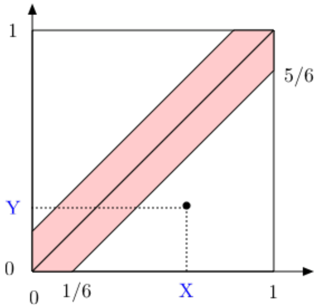
The complement is the sum of two rectangles. When you put them together, they form a square with sides $5/6$.

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

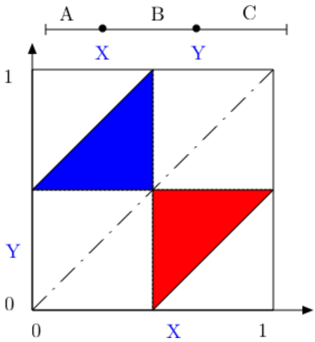
You break a stick at two points chosen independently uniformly at random.

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.

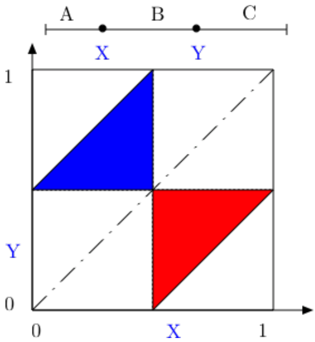
What is the probability you can make a triangle with the three pieces?



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?

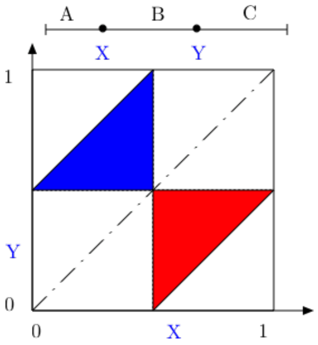


Let X, Y be the two break points along the $[0, 1]$ stick.

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?



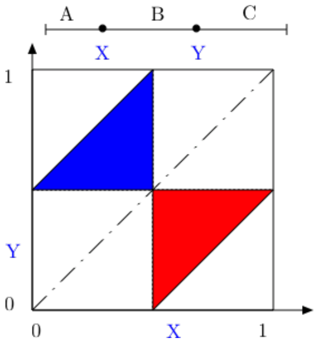
Let X, Y be the two break points along the $[0, 1]$ stick.

You can make 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.

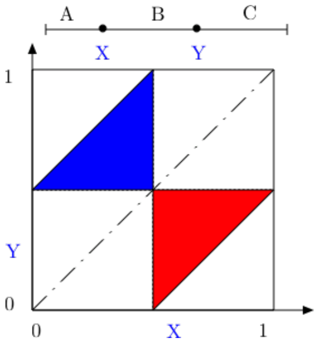
You can make a triangle if $A < B + C, B < A + C$, and $C < A + B$.

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

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.

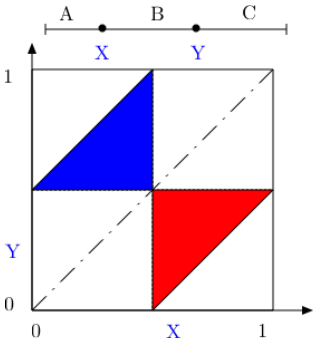
You can make a triangle if $A < B + C, B < A + C$, and $C < A + B$.

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

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.

You can make a triangle if $A < B + C, B < A + C$, and $C < A + B$.

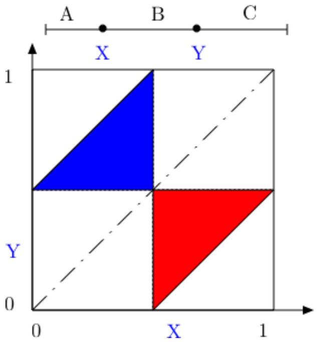
If $X < Y$, this means $X < 0.5, Y < X + 0.5, Y > 0.5$. This is the blue triangle.

If $X > Y$, we get the 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.

You can make a triangle if $A < B + C, B < A + C$, and $C < A + B$.

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

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

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

Maximum of Two Exponentials

Maximum of Two Exponentials

Let $X = \mathit{Expo}(\lambda)$ and $Y = \mathit{Expo}(\mu)$ 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 $X = Expo(\lambda)$ and $Y = Expo(\mu)$ 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

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

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,

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

Hence,

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

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,

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

Hence,

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

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,

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

Hence,

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

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

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

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 $SNR(dB) \approx 112 dB$.

Replacing Light Bulbs

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem:

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!} e^{-t}$.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!} e^{-t}$.

Proof:

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!} e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Let A be the event that a burns out during $[t, t+\varepsilon]$.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Let A be the event that a burns out during $[t, t+\varepsilon]$. Then,

$$Pr[X_{t+\varepsilon} = n] \quad \approx \quad Pr[X_t = n, A^c] + Pr[X_t = n-1, A]$$

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Let A be the event that a burns out during $[t, t+\varepsilon]$. Then,

$$\begin{aligned} Pr[X_{t+\varepsilon} = n] &\approx Pr[X_t = n, A^c] + Pr[X_t = n-1, A] \\ &= Pr[X_t = n]Pr[A^c] + Pr[X_t = n-1]Pr[A] \end{aligned}$$

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Let A be the event that a burns out during $[t, t+\varepsilon]$. Then,

$$\begin{aligned} Pr[X_{t+\varepsilon} = n] &\approx Pr[X_t = n, A^c] + Pr[X_t = n-1, A] \\ &= Pr[X_t = n]Pr[A^c] + Pr[X_t = n-1]Pr[A] \\ &\approx Pr[X_t = n](1-\varepsilon) + Pr[X_t = n-1]\varepsilon. \end{aligned}$$

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Let A be the event that a burns out during $[t, t+\varepsilon]$. Then,

$$\begin{aligned} Pr[X_{t+\varepsilon} = n] &\approx Pr[X_t = n, A^c] + Pr[X_t = n-1, A] \\ &= Pr[X_t = n]Pr[A^c] + Pr[X_t = n-1]Pr[A] \\ &\approx Pr[X_t = n](1-\varepsilon) + Pr[X_t = n-1]\varepsilon. \end{aligned}$$

Hence, $g(n,t) := Pr[X_t = n]$

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Let A be the event that a burns out during $[t, t+\varepsilon]$. Then,

$$\begin{aligned} Pr[X_{t+\varepsilon} = n] &\approx Pr[X_t = n, A^c] + Pr[X_t = n-1, A] \\ &= Pr[X_t = n]Pr[A^c] + Pr[X_t = n-1]Pr[A] \\ &\approx Pr[X_t = n](1-\varepsilon) + Pr[X_t = n-1]\varepsilon. \end{aligned}$$

Hence, $g(n,t) := Pr[X_t = n]$ is such that

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: We see how X_t increases over the next $\varepsilon \ll 1$ time units.

Let A be the event that a burns out during $[t, t+\varepsilon]$. Then,

$$\begin{aligned} Pr[X_{t+\varepsilon} = n] &\approx Pr[X_t = n, A^c] + Pr[X_t = n-1, A] \\ &= Pr[X_t = n]Pr[A^c] + Pr[X_t = n-1]Pr[A] \\ &\approx Pr[X_t = n](1-\varepsilon) + Pr[X_t = n-1]\varepsilon. \end{aligned}$$

Hence, $g(n,t) := Pr[X_t = n]$ is such that

$$g(n, t+\varepsilon) \approx g(n,t) - g(n,t)\varepsilon + g(n-1,t)\varepsilon.$$

Replacing Light Bulbs

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem:

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!} e^{-t}$.

Proof: (continued)

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: (continued) We saw that

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!} e^{-t}$.

Proof: (continued) We saw that

$$g(n, t+\varepsilon) \approx g(n,t) - g(n,t)\varepsilon + g(n-1,t)\varepsilon.$$

Replacing Light Bulbs

Say that light bulbs have i.i.d. *Expo*(1) lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: (continued) We saw that

$$g(n, t+\varepsilon) \approx g(n,t) - g(n,t)\varepsilon + g(n-1,t)\varepsilon.$$

Subtracting $g(n,t)$, dividing by ε, and letting $\varepsilon \to 0$,

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!} e^{-t}$.

Proof: (continued) We saw that

$$g(n, t + \varepsilon) \approx g(n, t) - g(n, t)\varepsilon + g(n-1, t)\varepsilon.$$

Subtracting $g(n, t)$, dividing by ε, and letting $\varepsilon \to 0$, one gets

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: (continued) We saw that

$$g(n, t+\varepsilon) \approx g(n,t) - g(n,t)\varepsilon + g(n-1,t)\varepsilon.$$

Subtracting $g(n,t)$, dividing by ε, and letting $\varepsilon \to 0$, one gets

$$g'(n,t) = -g(n,t) + g(n-1,t).$$

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: (continued) We saw that

$$g(n, t+\varepsilon) \approx g(n,t) - g(n,t)\varepsilon + g(n-1,t)\varepsilon.$$

Subtracting $g(n,t)$, dividing by ε, and letting $\varepsilon \to 0$, one gets

$$g'(n,t) = -g(n,t) + g(n-1,t).$$

You can check that these equations are solved by $g(n,t) = \frac{t^n}{n!}e^{-t}$.

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!}e^{-t}$.

Proof: (continued) We saw that

$$g(n, t+\varepsilon) \approx g(n,t) - g(n,t)\varepsilon + g(n-1,t)\varepsilon.$$

Subtracting $g(n,t)$, dividing by ε, and letting $\varepsilon \to 0$, one gets

$$g'(n,t) = -g(n,t) + g(n-1,t).$$

You can check that these equations are solved by $g(n,t) = \frac{t^n}{n!}e^{-t}$.
Indeed, then

$$g'(n,t) = \frac{t^{n-1}}{(n-1)!}e^{-t} - g(n,t)$$

Replacing Light Bulbs

Say that light bulbs have i.i.d. $Expo(1)$ lifetimes.

We turn a light on, and replace it as soon as it burns out.

How many light bulbs do we need to replace in t units of time?

Theorem: The number X_t of replaced light bulbs is $P(t)$.

That is, $Pr[X_t = n] = \frac{t^n}{n!} e^{-t}$.

Proof: (continued) We saw that

$$g(n, t+\varepsilon) \approx g(n,t) - g(n,t)\varepsilon + g(n-1,t)\varepsilon.$$

Subtracting $g(n,t)$, dividing by ε, and letting $\varepsilon \to 0$, one gets

$$g'(n,t) = -g(n,t) + g(n-1,t).$$

You can check that these equations are solved by $g(n,t) = \frac{t^n}{n!} e^{-t}$. Indeed, then

$$\begin{aligned} g'(n,t) &= \frac{t^{n-1}}{(n-1)!} e^{-t} - g(n,t) \\ &= g(n-1,t) - g(n,t). \end{aligned}$$

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

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

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 3

Summary

Continuous Probability 3

Summary

Continuous Probability 3

- Continuous RVs are essentially the same as discrete RVs

Summary

Continuous Probability 3

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

Summary

Continuous Probability 3

- 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 3

- 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 3

- 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.