Calculus Review

$$\frac{d(e^{cx})}{dx} = ce^{cx}.$$

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have!

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $\qquad e = (1 + 1/n)^n$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $\quad e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y$

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx}$

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1+1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1+1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Slope of $g(\cdot)$ times slope of $f(\cdot)$ at appropriate values.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Slope of $g(\cdot)$ times slope of $f(\cdot)$ at appropriate values.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $\quad e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Slope of $g(\cdot)$ times slope of $f(\cdot)$ at appropriate values.

Product Rule:

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Slope of $g(\cdot)$ times slope of $f(\cdot)$ at appropriate values.

Product Rule:

$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Slope of $g(\cdot)$ times slope of $f(\cdot)$ at appropriate values.

Product Rule:

$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$.

$d(uv) = udv + vdu$

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Slope of $g(\cdot)$ times slope of $f(\cdot)$ at appropriate values.

Product Rule:

$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$.

$d(uv) = udv + vdu$

Cuz:$d(uv) = uv - (u + du)(v + dv) = udv + vdu + dudv$.

Calculus Review

$\frac{d(e^{cx})}{dx} = ce^{cx}$.

Grows proportional to what you have! $e = (1 + 1/n)^n$.

$\frac{d(x^2)}{dx} = 2x$.

$\frac{(x+\delta)^2 - x^2}{\delta} = \frac{2x\delta + \delta^2}{\delta} = 2x + \delta$.

$\int x dx = \frac{x^2}{2} + c$.

Fundamental Theorem. or Triangle: width x, height x has area $\frac{x^2}{2}$.

$\frac{d(\ln x)}{dx} = \frac{1}{x}$

$x = e^y \implies 1 = e^y \frac{dy}{dx} \implies \frac{dy}{dx} = \frac{1}{e^y} = \frac{1}{x}$.

Flipping x and y axis, flips slope and function and argument.

Chain Rule: $\frac{d(f(g(x))}{dx} = f'(g(x))g'(x)dx$

Slope of $g(\cdot)$ times slope of $f(\cdot)$ at appropriate values.

Product Rule:

$(f(x)g(x))' = f'(x)g(x) + f(x)g'(x)$.

$d(uv) = udv + vdu$

Cuz:$d(uv) = uv - (u + du)(v + dv) = udv + vdu + dudv$.

Integration by Parts: $\int udv = uv - \int vdu$.

Summary

Continuous Probability 1

1. pdf:

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta.$
2. CDF:

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$:

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}$;

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\};$

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.
5. Target:

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.
5. Target: $f_X(x) = 2x \cdot 1\{0 \leq x \leq 1\}$;

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.
5. Target: $f_X(x) = 2x \cdot 1\{0 \leq x \leq 1\}; F_X(x) = x^2$ for $0 \leq x \leq 1$.

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.
5. Target: $f_X(x) = 2x \cdot 1\{0 \leq x \leq 1\}; F_X(x) = x^2$ for $0 \leq x \leq 1$.
6. Joint pdf: $Pr[X \in (x, x+\delta), Y = (y, y+\delta)) = f_{X,Y}(x,y)\delta^2$.

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.
5. Target: $f_X(x) = 2x \cdot 1\{0 \leq x \leq 1\}; F_X(x) = x^2$ for $0 \leq x \leq 1$.
6. Joint pdf: $Pr[X \in (x, x+\delta), Y = (y, y+\delta)) = f_{X,Y}(x,y)\delta^2$.
   6.1 Conditional Distribution: $f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.
5. Target: $f_X(x) = 2x \cdot 1\{0 \leq x \leq 1\}; F_X(x) = x^2$ for $0 \leq x \leq 1$.
6. Joint pdf: $Pr[X \in (x, x+\delta), Y = (y, y+\delta)) = f_{X,Y}(x,y)\delta^2$.
   6.1 Conditional Distribution: $f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.
   6.2 Independence: $f_{X|Y}(x,y) = f_X(x)$

Summary

Continuous Probability 1

1. pdf: $Pr[X \in (x, x+\delta]] \approx f_X(x)\delta$.
2. CDF: $Pr[X \leq x] = F_X(x) = lim_{\delta \to 0} \sum_i f_X(x_i)\delta = \int_{-\infty}^{x} f_X(y)dy$.
3. $X \sim U[a,b]$: $f_X(x) = \frac{1}{b-a} 1\{a \leq x \leq b\}; F_X(x) = \frac{x-a}{b-a}$ for $a \leq x \leq b$.
4. $X \sim Expo(\lambda)$:
   $f_X(x) = \lambda \exp\{-\lambda x\} 1\{x \geq 0\}; F_X(x) = 1 - \exp\{-\lambda x\}$ for $x \leq 0$.
5. Target: $f_X(x) = 2x \cdot 1\{0 \leq x \leq 1\}; F_X(x) = x^2$ for $0 \leq x \leq 1$.
6. Joint pdf: $Pr[X \in (x, x+\delta), Y = (y, y+\delta)) = f_{X,Y}(x,y)\delta^2$.
   6.1 Conditional Distribution: $f_{X|Y}(x,y) = \frac{f_{X,Y}(x,y)}{f_Y(y)}$.
   6.2 Independence: $f_{X|Y}(x,y) = f_X(x)$

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.
Event $Y > y$ is event $X > y/2$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.
Event $Y > y$ is event $X > y/2$.

(E) $Y = 2X$, $f_Y(y) = \frac{1}{2}f(y/2)$.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.
Event $Y > y$ is event $X > y/2$.

(E) $Y = 2X$, $f_Y(y) = \frac{1}{2}f(y/2)$.
Spreads out density of X over twice the range.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.
Event $Y > y$ is event $X > y/2$.

(E) $Y = 2X$, $f_Y(y) = \frac{1}{2}f(y/2)$.
Spreads out density of X over twice the range.
Chain rule from (D).

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.
Event $Y > y$ is event $X > y/2$.

(E) $Y = 2X$, $f_Y(y) = \frac{1}{2}f(y/2)$.
Spreads out density of X over twice the range.
Chain rule from (D).

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.
Event $Y > y$ is event $X > y/2$.

(E) $Y = 2X$, $f_Y(y) = \frac{1}{2}f(y/2)$.
Spreads out density of X over twice the range.
Chain rule from (D).

(A), (B), (D) think events, (E) think event and density.

Poll

What is true?
X has CDF $F(x)$ and PDF $f(x)$.

(A) $Pr[X > t] = 1 - Pr[X \leq t]$.
Event $X > t$ is the event that X is not $\leq t$.

(B) $S(t) = Pr[X > t] = 1 - F(t)$.
Definition of CDF.

(C) $Y = 2X$, $f_Y(y) = 2f(y)$.
False. Confuses density of outcome with value oof outcome.

(D) $Y = 2X$, $F_Y(y) = F(y/2)$.
Event $Y > y$ is event $X > y/2$.

(E) $Y = 2X$, $f_Y(y) = \frac{1}{2}f(y/2)$.
Spreads out density of X over twice the range.
Chain rule from (D).

(A), (B), (D) think events, (E) think event and density.

(C) confuses probability density of outcome with value of outcome.

Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: "outcome" is real number.

Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: "outcome" is real number.
Probability: Events is interval.

Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: “outcome” is real number.
Probability: Events is interval.
Density: $Pr[X \in [x, x+dx]] = f(x)dx$

Discrete/Continuous

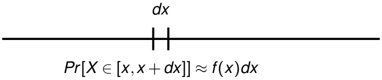
Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: "outcome" is real number.
Probability: Events is interval.
Density: $Pr[X \in [x, x+dx]] = f(x)dx$

Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

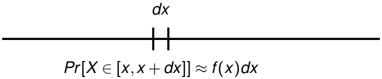
Continuous: “outcome” is real number.
Probability: Events is interval.
Density: $Pr[X \in [x, x+dx]] = f(x)dx$



Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: “outcome” is real number.
Probability: Events is interval.
Density: $Pr[X \in [x, x+dx]] = f(x)dx$

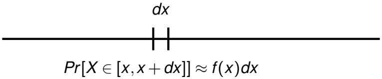


Joint Continuous in d variables: “outcome” is $\in R^d$.
Probability: Events is block.

Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: "outcome" is real number.
  Probability: Events is interval.
   Density: $Pr[X \in [x, x+dx]] = f(x)dx$

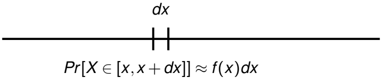


Joint Continuous in d variables: "outcome" is $\in R^d$.
   Probability: Events is block.
   Density: $Pr[(X,Y) \in ([x, x+dx],[y, y+dx])] = f(x,y)dxdy$

Discrete/Continuous

Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: "outcome" is real number.
Probability: Events is interval.
Density: $Pr[X \in [x, x+dx]] = f(x)dx$

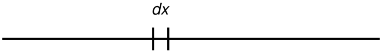


Joint Continuous in d variables: "outcome" is $\in R^d$.
Probability: Events is block.
Density: $Pr[(X,Y) \in ([x, x+dx], [y, y+dx])] = f(x,y)dxdy$

Discrete/Continuous

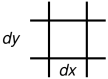
Discrete: Probability of outcome $\rightarrow$ random variables, events.

Continuous: "outcome" is real number.
Probability: Events is interval.
Density: $Pr[X \in [x, x+dx]] = f(x)dx$



$Pr[X \in [x, x+dx]] \approx f(x)dx$

Joint Continuous in d variables: "outcome" is $\in R^d$.
Probability: Events is block.
Density: $Pr[(X, Y) \in ([x, x+dx], [y, y+dx])] = f(x, y)dxdy$



$Pr[(X, Y) \in ([x, x+dx], [y, y+dy])] \approx f(x, y)dxdy$

Probability

Probability

Probability!

Probability

Probability!
Challenges us.

Probability

Probability!
Challenges us. But really neat.

Probability

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

Probability

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

Probability

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

Probability

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

Probability

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

Probability

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

Probability

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

Probability

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

Probability

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

Sample Space:Ω, $Pr[\omega]$.

Probability

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

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

Probability

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

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

Random Variable: X, Range is reals.
Event: $A = [a,b]$, $Pr[X \in A]$,

Probability

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

Sample Space: Ω, $Pr[\omega]$.
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, Range is reals.
Event: $A = [a, b]$, $Pr[X \in A]$,
CDF: $F(x) = Pr[X \le x]$.
PDF: $f(x) = \frac{dF(x)}{dx}$.
$\int_{-\infty}^{\infty} f(x) = 1$.

Probability

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

Sample Space: Ω, $Pr[\omega]$.
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, Range is reals.
Event: $A = [a, b]$, $Pr[X \in A]$,
CDF: $F(x) = Pr[X \leq x]$.
PDF: $f(x) = \frac{dF(x)}{dx}$.
$\int_{-\infty}^{\infty} f(x) = 1$.

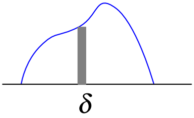
Probability

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

Sample Space: Ω, $Pr[\omega]$.
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, Range is reals.
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 Probability.

Probability Rules are all good.

Conditional Probability.

Events: A, B

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

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 Probability.

Events: A, B

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

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

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

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

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

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"] $= Pr[HH] + Pr[TH]$
B is First coin heads.

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

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

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

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

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

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

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

Probability Rules are all good.

Conditional Probability.

Events: A, B

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

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

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

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"] $= Pr[HH] + Pr[TH]$
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 Probability.

Events: A, B

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

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

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

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

Total Probability Rule: $Pr[A] = Pr[A \cap B] + Pr[A \cap \overline{B}]$
$Pr[\text{“Second Heads”}] = Pr[HH] + Pr[TH]$
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.

Conditional density.

Conditional Density: $f_{X|Y}(x,y)$.

Conditional density.

Conditional Density: $f_{X|Y}(x,y)$.

Conditional Probability: $Pr[X \in A | Y \in B] = \frac{Pr[X \in A, Y \in B]}{Pr[Y \in B]}$

Conditional density.

Conditional Density: $f_{X|Y}(x,y)$.

Conditional Probability: $Pr[X \in A | Y \in B] = \frac{Pr[X \in A, Y \in B]}{Pr[Y \in B]}$

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

Conditional density.

Conditional Density: $f_{X|Y}(x,y)$.

Conditional Probability: $Pr[X \in A | Y \in B] = \frac{Pr[X \in A, Y \in B]}{Pr[Y \in B]}$

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

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

Conditional density.

Conditional Density: $f_{X|Y}(x,y)$.

Conditional Probability: $Pr[X \in A | Y \in B] = \frac{Pr[X \in A, Y \in B]}{Pr[Y \in B]}$

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

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

Corollary: For independent random variables, $f_{X|Y}(x,y) = f_X(x)$.

Expo(λ)

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

$Expo(\lambda)$

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

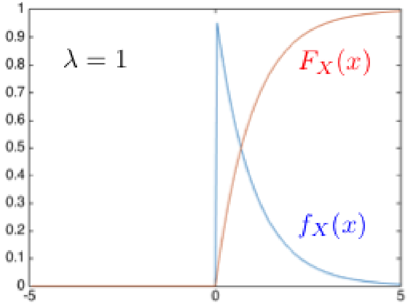
$$f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$$

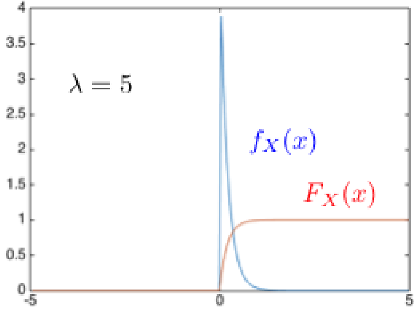
$Expo(\lambda)$

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

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

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



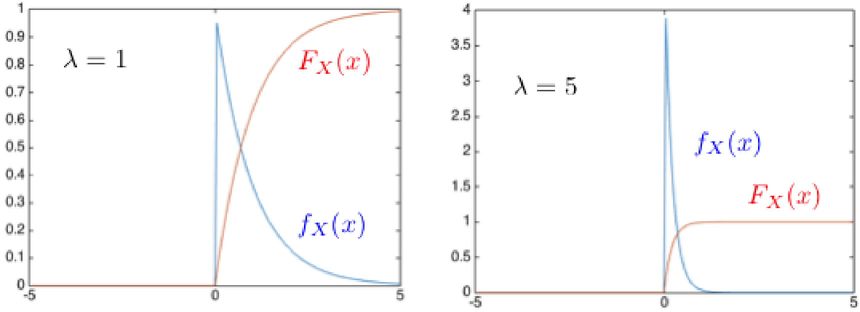


$Expo(\lambda)$

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 $s, t > 0$,

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

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 as 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 as 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 as 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 as good as new.'

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

$$Pr[Y > t] \quad =$$

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

Replace b by $b-a$, use $X = U[0,1]$, 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

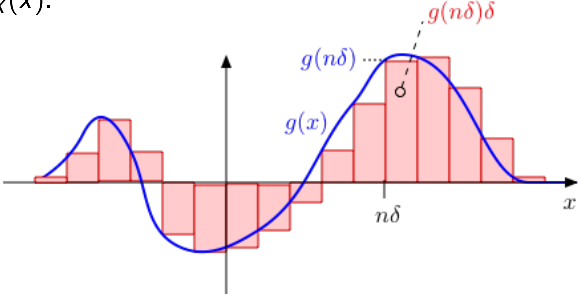
Definition: The **expectation** of a random variable X with pdf $f(x)$ is *defined* as

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

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, $f_X(x) = 1\{0 \le x \le 1\}$. 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, $f_X(x) = 1\{0 \le x \le 1\}$. 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,

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

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,

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

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,

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

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 $f_X(x) = 2x 1\{0 \leq x \leq 1\}$. Thus,

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

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,

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

Recall the integration by parts formula:

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

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,

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

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,

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

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:

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

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:

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

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

Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**)

Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}.$$

Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}.$$

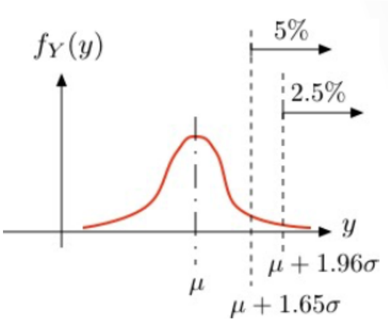
Standard normal has $\mu = 0$ **and** $\sigma = 1$**.**

Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}.$$

Standard normal has $\mu = 0$ **and** $\sigma = 1$**.**

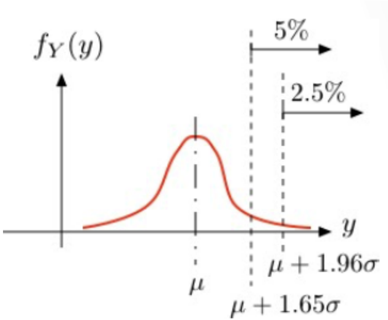


Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}.$$

Standard normal has $\mu = 0$ and $\sigma = 1$.



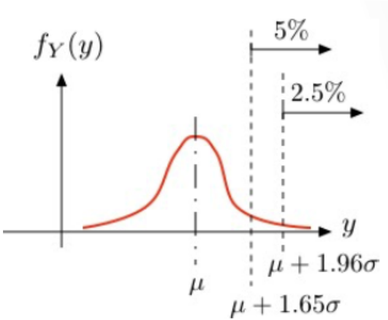
Note: $Pr[|Y - \mu| > 1.65\sigma] \approx 10\%$;

Normal (Gaussian) Distribution.

For any μ and σ, a **normal** (aka **Gaussian**) random variable Y, which we write as $Y = \mathcal{N}(\mu, \sigma^2)$, has pdf

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-(y-\mu)^2/2\sigma^2}.$$

Standard normal has $\mu = 0$ and $\sigma = 1$.



Note: $Pr[|Y - \mu| > 1.65\sigma] \approx 10\%$; $Pr[|Y - \mu| > 2\sigma] \approx 5\%$.

Scaling and Shifting and properties

Theorem Let $X = \mathcal{N}(0,1)$ and $Y = \mu + \sigma X$. Then

$$Y = \mathcal{N}(\mu, \sigma^2).$$

Scaling and Shifting and properties

Theorem Let $X = \mathcal{N}(0,1)$ and $Y = \mu + \sigma X$. Then

$$Y = \mathcal{N}(\mu, \sigma^2).$$

Theorem If $Y = \mathcal{N}(\mu, \sigma^2)$, then

$$E[Y] = \mu$$

Scaling and Shifting and properties

Theorem Let $X = \mathcal{N}(0,1)$ and $Y = \mu + \sigma X$. Then

$$Y = \mathcal{N}(\mu, \sigma^2).$$

Theorem If $Y = \mathcal{N}(\mu, \sigma^2)$, then

$$E[Y] = \mu \text{ and } var[Y] = \sigma^2.$$

Review: Law of Large Numbers.

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n} \sum X_i$ "tends to the mean."

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Each X_i, has $\mu = E(X_i)$ and $\sigma^2 = \text{var}(X_i)$.

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Each X_i, has $\mu = E(X_i)$ and $\sigma^2 = \text{var}(X_i)$.

Mean of A_n is μ, and variance is σ^2/n.

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Each X_i, has $\mu = E(X_i)$ and $\sigma^2 = \text{var}(X_i)$.

Mean of A_n is μ, and variance is σ^2/n.

Used Chebyshev.

$$Pr[|A_n - \mu| > \varepsilon] \leq$$

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Each X_i, has $\mu = E(X_i)$ and $\sigma^2 = \text{var}(X_i)$.

Mean of A_n is μ, and variance is σ^2/n.

Used Chebyshev.

$$Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2} =$$

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Each X_i, has $\mu = E(X_i)$ and $\sigma^2 = \text{var}(X_i)$.

Mean of A_n is μ, and variance is σ^2/n.

Used Chebyshev.

$$Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon}$$

Review: Law of Large Numbers.

Theorem: Independent identically distributed random variables, X_i,

$A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Each X_i, has $\mu = E(X_i)$ and $\sigma^2 = \text{var}(X_i)$.

Mean of A_n is μ, and variance is σ^2/n.

Used Chebyshev.

$$Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2} = \frac{\sigma^2}{n\varepsilon} \rightarrow 0.$$

Central Limit Theorem

Central Limit Theorem

Central Limit Theorem

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$.

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof:

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

$$E(S_n)$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

$$E(S_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu)$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

$$E(S_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu) = 0$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

$$E(S_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu) = 0$$

$$Var(S_n)$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

$$E(S_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu) = 0$$

$$Var(S_n) = \frac{1}{\sigma^2/n} Var(A_n)$$

Central Limit Theorem

Central Limit Theorem

Let $X_1, X_2, \ldots$ be i.i.d. with $E[X_1] = \mu$ and $var(X_1) = \sigma^2$. Define

$$S_n := \frac{A_n - \mu}{\sigma/\sqrt{n}} = \frac{X_1 + \cdots + X_n - n\mu}{\sigma\sqrt{n}}.$$

Then,

$$S_n \to \mathcal{N}(0,1), \text{as } n \to \infty.$$

That is,

$$Pr[S_n \leq \alpha] \to \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\alpha} e^{-x^2/2} dx.$$

Proof: See EE126.

Note:

$$E(S_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu) = 0$$
$$Var(S_n) = \frac{1}{\sigma^2/n} Var(A_n) = 1.$$

Confidence Intervals.

Recall: $A_n = \frac{1}{n} \sum X_i$, for X_i identical and independent.

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq$

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

$\frac{var A_n}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \leq \delta$ or $n \geq \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

$\frac{var A_n}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \leq \delta$ or $n \geq \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Central Limit Theorem:

$Pr[|A_n - \mu| > \varepsilon] \leq$

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

$\frac{var A_n}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \leq \delta$ or $n \geq \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Central Limit Theorem:

$$Pr[|A_n - \mu| > \varepsilon] \leq C\int_{x\geq\varepsilon}^{\infty} e^{-\frac{x^2}{2 var A_n}} \leq Ce^{-\frac{\varepsilon^2}{2 var A_n}}$$

for $\varepsilon > \sqrt{Var A_n}$ (C is roughly $2/\sqrt{2\pi}$)

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

$\frac{var A_n}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \leq \delta$ or $n \geq \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Central Limit Theorem:

$Pr[|A_n - \mu| > \varepsilon] \leq C\int_{x \geq \varepsilon}^{\infty} e^{-\frac{x^2}{2 var A_n}} \leq Ce^{-\frac{\varepsilon^2}{2 var A_n}}$

for $\varepsilon > \sqrt{Var A_n}$ (C is roughly $2/\sqrt{2\pi}$)

Implies to get confidence $1 - C\delta$ we need

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

$\frac{var A_n}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \leq \delta$ or $n \geq \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Central Limit Theorem:

$$Pr[|A_n - \mu| > \varepsilon] \leq C\int_{x\geq\varepsilon}^{\infty} e^{-\frac{x^2}{2 var A_n}} \leq Ce^{-\frac{\varepsilon^2}{2 var A_n}}$$

for $\varepsilon > \sqrt{Var A_n}$ (C is roughly $2/\sqrt{2\pi}$)

Implies to get confidence $1 - C\delta$ we need

$$e^{-\frac{\varepsilon^2}{2 var A}} \leq \delta$$

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

$\frac{var A_n}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \leq \delta$ or $n \geq \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Central Limit Theorem:

$Pr[|A_n - \mu| > \varepsilon] \leq C\int_{x \geq \varepsilon}^{\infty} e^{-\frac{x^2}{2 var A_n}} \leq Ce^{-\frac{\varepsilon^2}{2 var A_n}}$

for $\varepsilon > \sqrt{Var A_n}$ (C is roughly $2/\sqrt{2\pi}$)

Implies to get confidence $1 - C\delta$ we need

$e^{-\frac{\varepsilon^2}{2 varA}} \leq \delta \implies \quad -\frac{n\varepsilon^2}{2\sigma^2} \leq \log \delta$

Confidence Intervals.

Recall: $A_n = \frac{1}{n}\sum X_i$, for X_i identical and independent.

For $\mu = E(X_i)$ and variance σ^2. Mean of A_n is μ, and variance is σ^2/n.

Recall Chebyshev: $Pr[|A_n - \mu| > \varepsilon] \leq \frac{var[A_n]}{\varepsilon^2}$

Implies to get confidence $1-\delta$ we need

$\frac{var A_n}{\varepsilon^2} = \frac{1}{n}\frac{\sigma^2}{\varepsilon^2} \leq \delta$ or $n \geq \frac{\sigma^2}{\varepsilon^2}\frac{1}{\delta}$

Central Limit Theorem:

$Pr[|A_n - \mu| > \varepsilon] \leq C\int_{x\geq\varepsilon}^{\infty} e^{-\frac{x^2}{2 var A_n}} \leq Ce^{-\frac{\varepsilon^2}{2 var A_n}}$

for $\varepsilon > \sqrt{Var A_n}$ (C is roughly $2/\sqrt{2\pi}$)

Implies to get confidence $1 - C\delta$ we need

$e^{-\frac{\varepsilon^2}{2 var A}} \leq \delta \implies \quad -\frac{n\varepsilon^2}{2\sigma^2} \leq \log\delta \implies n \geq \frac{2\sigma^2}{\varepsilon^2}\log\frac{1}{\delta}$.

Examples: Meeting at a Restaurant

Examples: Meeting at a Restaurant

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

Examples: 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.

Examples: 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.

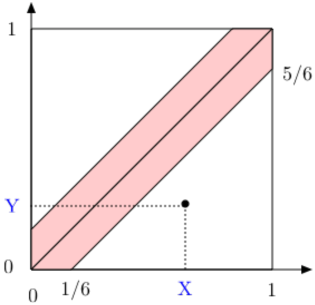
What is the probability they meet?

Examples: 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?

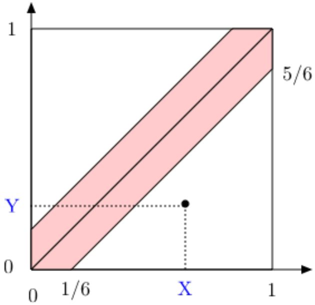


Examples: Meeting at a Restaurant

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?



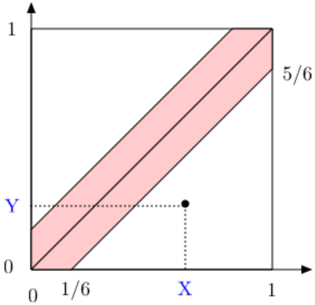
Here, (X, Y) are the times when the friends reach the restaurant.

Examples: 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.

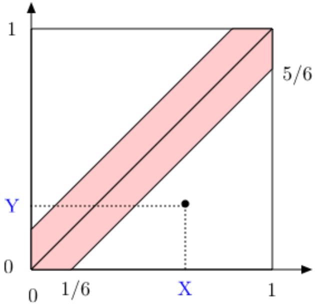
The shaded area are the pairs where $|X - Y| < 1/6$,

Examples: 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.

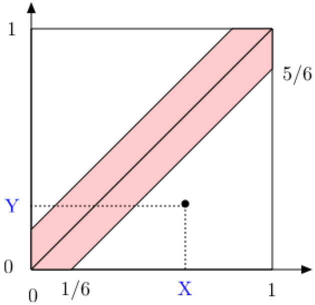
The shaded area are the pairs where $|X - Y| < 1/6$, i.e., such that they meet.

Examples: 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.

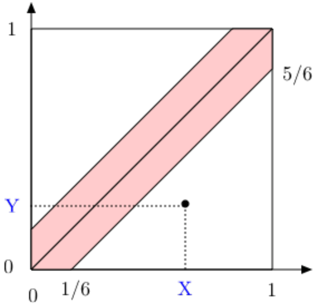
The complement is the sum of two rectangles.

Examples: 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.

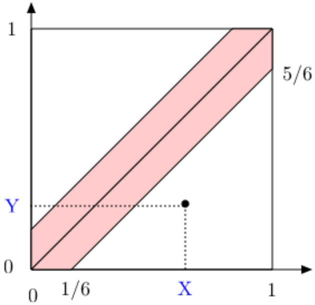
The complement is the sum of two rectangles. When you put them together, they form a square

Examples: 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.

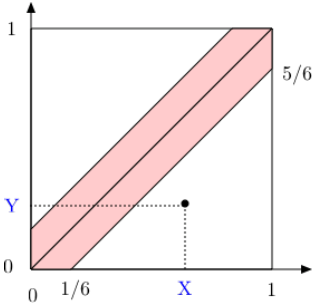
The complement is the sum of two rectangles. When you put them together, they form a square with sides $5/6$.

Examples: 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 $5/6$.

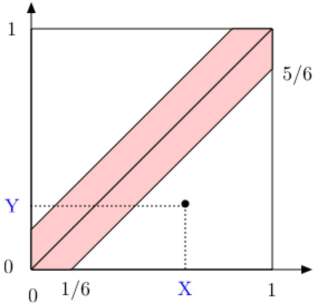
Thus, $Pr[\text{meet}] = 1 - (\frac{5}{6})^2 =$

Examples: 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

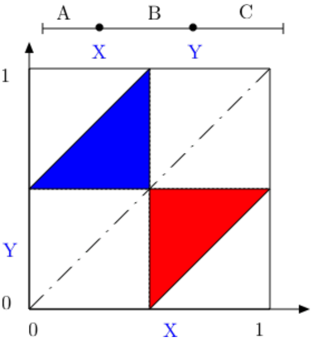
You break a stick at two points chosen independently uniformly at random.

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.

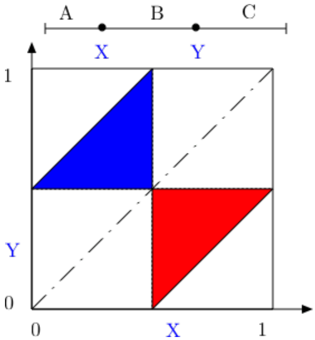
What is the probability you can make a triangle with the three pieces?



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?

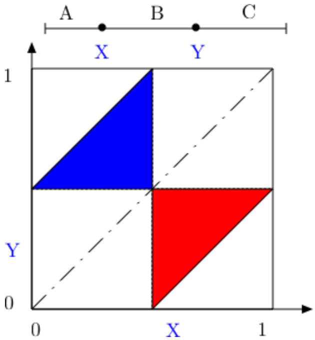


Let X, Y be the two break points along the $[0, 1]$ stick.

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?



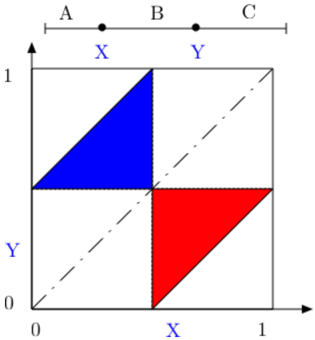
Let X, Y be the two break points along the $[0,1]$ stick.

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.

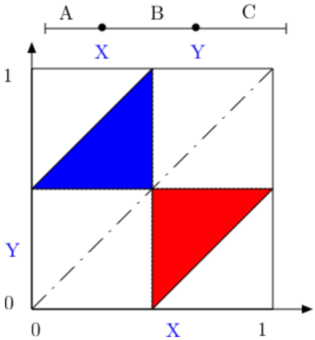
A triangle if
$A < B + C, B < A + C$, and $C < A + B$.

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.

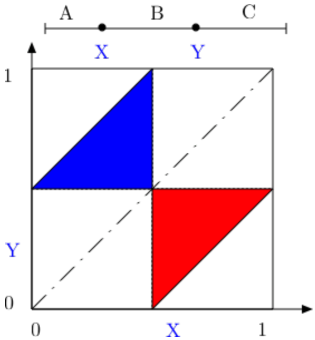
A triangle if
$A < B + C, B < A + C$, and $C < A + B$.

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.

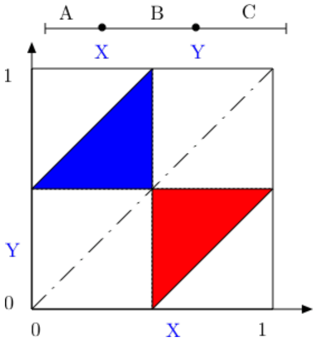
A triangle if $A < B + C, B < A + C$, and $C < A + B$.

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.

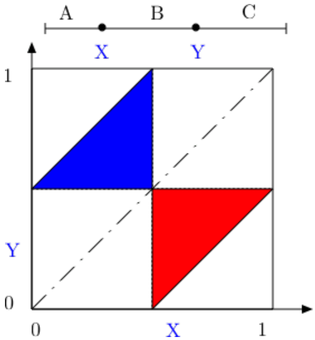
A triangle if
$A < B + C, B < A + C$, and $C < A + B$.

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 $C < A + B$.

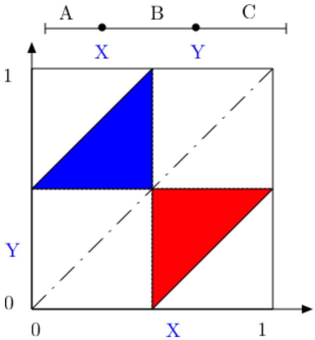
If $X < Y$, this means
$X < 0.5$, $Y < X + .5$, $Y > 0.5$.
This is the blue triangle.

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 $C < A + B$.

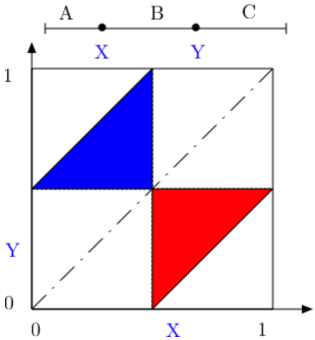
If $X < Y$, this means
$X < 0.5$, $Y < X + .5$, $Y > 0.5$.
This is the blue triangle.

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

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

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,

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

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

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

Minimum of n i.i.d. Exponentials.

Let $X_1, \ldots, X_n$ be i.i.d. $Expo(1)$.

Minimum of n i.i.d. Exponentials.

Let $X_1, \ldots, X_n$ be i.i.d. *Expo*(1). Define $Z = \min\{X_1, X_2, \ldots, X_n\}$.

Minimum of n i.i.d. Exponentials.

Let $X_1,\ldots,X_n$ be i.i.d. *Expo*(1). Define $Z = \min\{X_1, X_2, \ldots, X_n\}$.

What is true?

Minimum of n i.i.d. Exponentials.

Let $X_1, \ldots, X_n$ be i.i.d. $Expo(1)$. Define $Z = \min\{X_1, X_2, \ldots, X_n\}$.

What is true?

(A) Z is exponential.
(B) Parameter is n.
(C) $lim_{N \to \infty}(1 - n/N)^N \to e^{-n}$
(D) $E[Z] = 1/n$.

Minimum of n i.i.d. Exponentials.

Let $X_1, \ldots, X_n$ be i.i.d. $Expo(1)$. Define $Z = \min\{X_1, X_2, \ldots, X_n\}$.

What is true?

(A) Z is exponential.
(B) Parameter is n.
(C) $lim_{N \to \infty}(1 - n/N)^N \to e^{-n}$
(D) $E[Z] = 1/n$.

(C) is an argument for (A), (B) and (D).

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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}.$$

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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}. \qquad Y_i \sim Expo(1).$$

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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}. \qquad Y_i \sim Expo(1).$$

From 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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}. \qquad Y_i \sim Expo(1).$$

From 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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}. \qquad Y_i \sim \mathit{Expo}(1).$$

From 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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}. \qquad Y_i \sim Expo(1).$$

From 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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}. \qquad Y_i \sim \mathit{Expo}(1).$$

From 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:

$$Z = \min\{X_1, \ldots, X_n\} + \max Y_1, \ldots, Y_{n-1}. \qquad Y_i \sim Expo(1).$$

From 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 - 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 - closest multiple of 2^{-n} to X. 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 - closest multiple of 2^{-n} to X. 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 - closest multiple of 2^{-n} to X. 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 - closest multiple of 2^{-n} to X. 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 - closest multiple of 2^{-n} to X. 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 $[-a, a = 2^{-(n)}]$.

Quantization Noise

In digital video and audio, one represents a continuous value by a finite number of bits.

This introduces an error perceived as noise: the quantization noise. What is the power of that noise?

Model: $X = U[0,1]$ is the continuous value. Y - closest multiple of 2^{-n} to X. 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 $[-a, a = 2^{-(n)}]$.

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 - closest multiple of 2^{-n} to X. 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 $[-a, a = 2^{-(n)}]$.

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 - closest multiple of 2^{-n} to X. 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 $[-a, a = 2^{-(n)}]$.

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 - closest multiple of 2^{-n} to X. 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 $[-a, a = 2^{-(n)}]$.

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

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

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

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

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? $\frac{n}{6}$.

Summary

Continuous Probability

Summary

Continuous Probability

Summary

Continuous Probability

- Continuous RVs are similar to discrete RVs

Summary

Continuous Probability

- Continuous RVs are similar to discrete RVs
- Think that $X \in [x, x+\varepsilon]$ with probability $f_X(x)\varepsilon$

Summary

Continuous Probability

- Continuous RVs are similar to discrete RVs
- Think that $X \in [x, x+\varepsilon]$ with probability $f_X(x)\varepsilon$
- Sums become integrals, ...

Summary

Continuous Probability

- Continuous RVs are similar to discrete RVs
- Think that $X \in [x, x+\varepsilon]$ with probability $f_X(x)\varepsilon$
- Sums become integrals, ...
- The exponential distribution is magical:

Summary

Continuous Probability

- Continuous RVs are similar to discrete RVs
- Think that $X \in [x, x + \varepsilon]$ with probability $f_X(x)\varepsilon$
- Sums become integrals, ...
- The exponential distribution is magical: memoryless.