CS70: Jean Walrand: Lecture 36.

Gaussian and CLT

CS70: Jean Walrand: Lecture 36.

Gaussian and CLT

Warning:

CS70: Jean Walrand: Lecture 36.

Gaussian and CLT

Warning: This lecture is also rated R.

CS70: Jean Walrand: Lecture 36.

Gaussian and CLT

Warning: This lecture is also rated R.

1. Review of continuous probability
2. Motivation for Gaussian
3. Gaussian
4. CLT

Review of Continuous Probability

Ω is continuous space.

Review of Continuous Probability

Ω is continuous space.
Probability of any outcome is 0.

Review of Continuous Probability

Ω is continuous space.
Probability of any outcome is 0.
Work with events.

Review of Continuous Probability

Ω is continuous space.
Probability of any outcome is 0.
Work with events.
Example: James Bond lands on position uniformly $[0, 1000]$.

Review of Continuous Probability

Ω is continuous space.
Probability of any outcome is 0.
Work with events.
Example: James Bond lands on position uniformly $[0, 1000]$.
Probability lands in an interval $[a, b] \subseteq [0, 1000]$ is

Review of Continuous Probability

Ω is continuous space.
Probability of any outcome is 0.
Work with events.
Example: James Bond lands on position uniformly $[0, 1000]$.
Probability lands in an interval $[a, b] \subseteq [0, 1000]$ is

$$\frac{b-a}{1000}.$$

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$

   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$

   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$ or $f_X(x) = \frac{d(F_X(x))}{dx}$.

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$ or $f_X(x) = \frac{d(F_X(x))}{dx}$.
   Probability Density Function (pdf).

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$ or $f_X(x) = \frac{d(F_X(x))}{dx}$.
   Probability Density Function (pdf).
   $Pr[a < X \leq b] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$ or $f_X(x) = \frac{d(F_X(x))}{dx}$.
   Probability Density Function (pdf).
   $Pr[a < X \leq b] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$
   2.1 $f_X(x) \geq 0$ for all $x \in \Re$.

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$ or $f_X(x) = \frac{d(F_X(x))}{dx}$.
   Probability Density Function (pdf).
   $Pr[a < X \leq b] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$
   2.1 $f_X(x) \geq 0$ for all $x \in \Re$.
   2.2 $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Random Variables

Continuous random variable X, specified by

1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$ or $f_X(x) = \frac{d(F_X(x))}{dx}$.
   Probability Density Function (pdf).
   $Pr[a < X \leq b] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$
   2.1 $f_X(x) \geq 0$ for all $x \in \Re$.
   2.2 $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Recall that $Pr[X \in (x, x+\delta)] \approx f_X(x)\delta$.

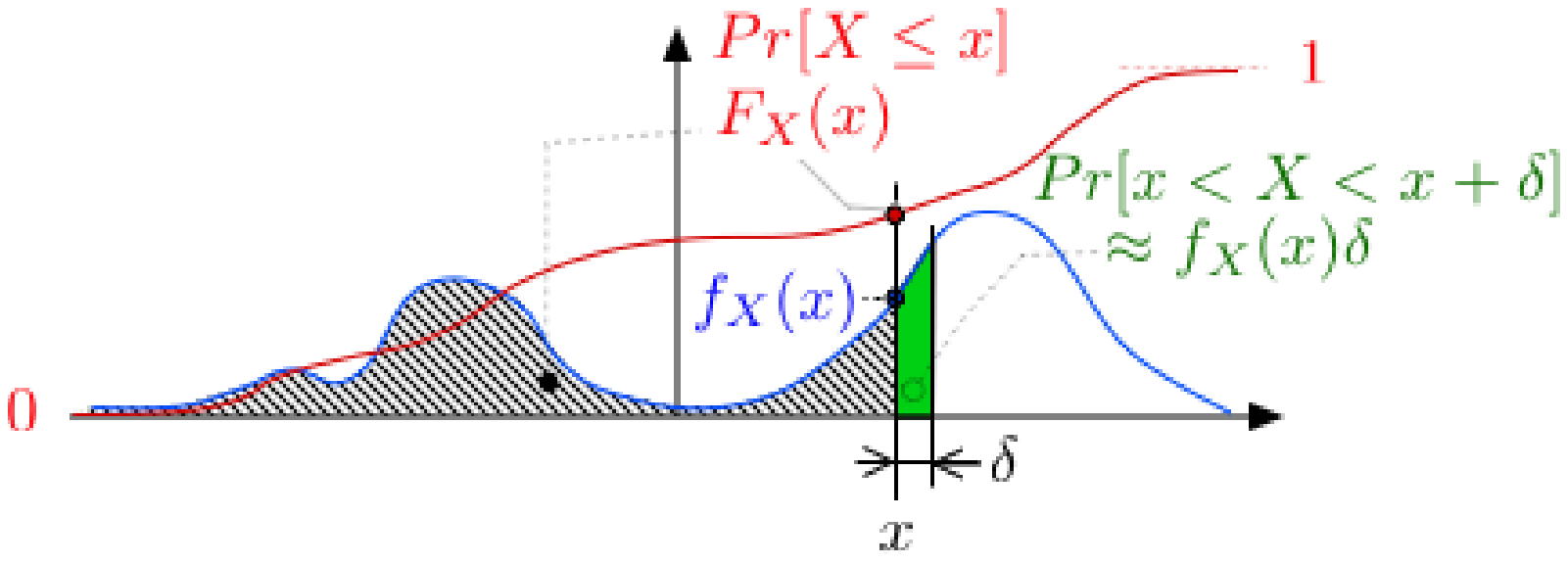
Random Variables

Continuous random variable X, specified by

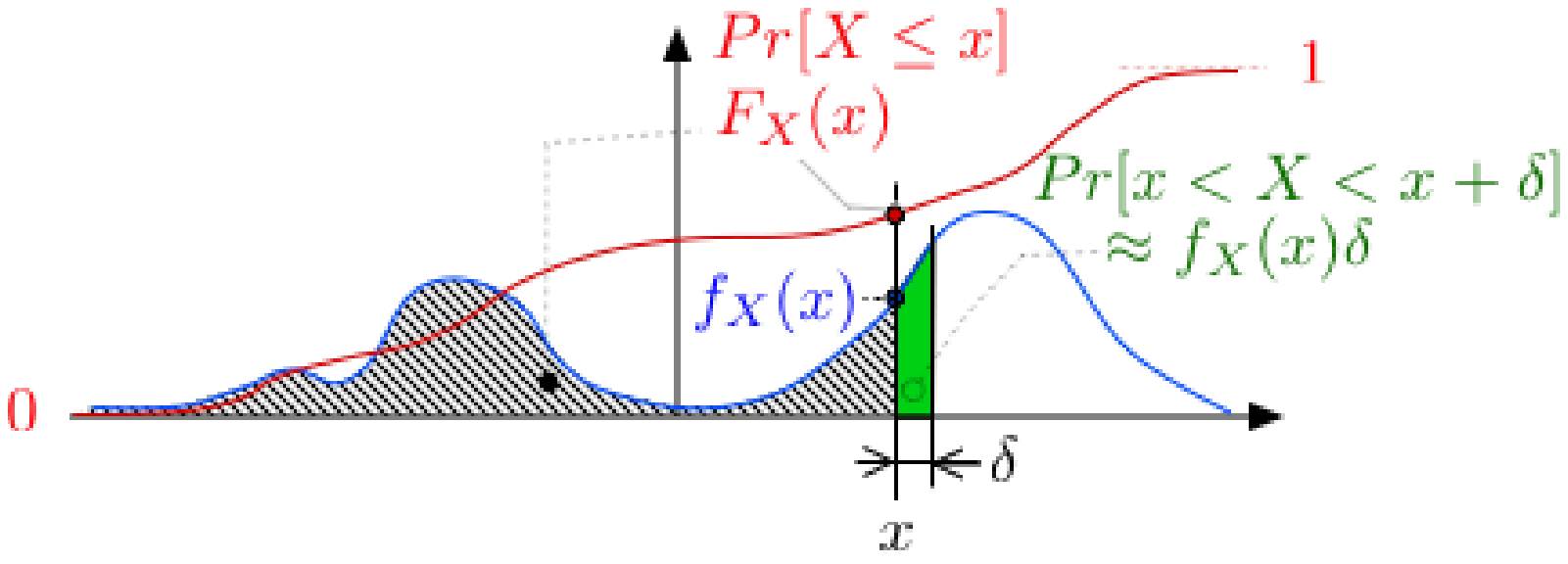
1. $F_X(x) = Pr[X \leq x]$ for all x.
   Cumulative Distribution Function (cdf).
   $Pr[a < X \leq b] = F_X(b) - F_X(a)$
   1.1 $0 \leq F_X(x) \leq 1$ for all $x \in \Re$.
   1.2 $F_X(x) \leq F_X(y)$ if $x \leq y$.
2. Or $f_X(x)$, where $F_X(x) = \int_{-\infty}^{x} f_X(y)dy$ or $f_X(x) = \frac{d(F_X(x))}{dx}$.
   Probability Density Function (pdf).
   $Pr[a < X \leq b] = \int_a^b f_X(x)dx = F_X(b) - F_X(a)$
   2.1 $f_X(x) \geq 0$ for all $x \in \Re$.
   2.2 $\int_{-\infty}^{\infty} f_X(x)dx = 1$.

Recall that $Pr[X \in (x, x+\delta)] \approx f_X(x)\delta$. Think of X taking discrete values $n\delta$ for $n = \ldots, -2, -1, 0, 1, 2, \ldots$ with $Pr[X = n\delta] = f_X(n\delta)\delta$.

A Picture

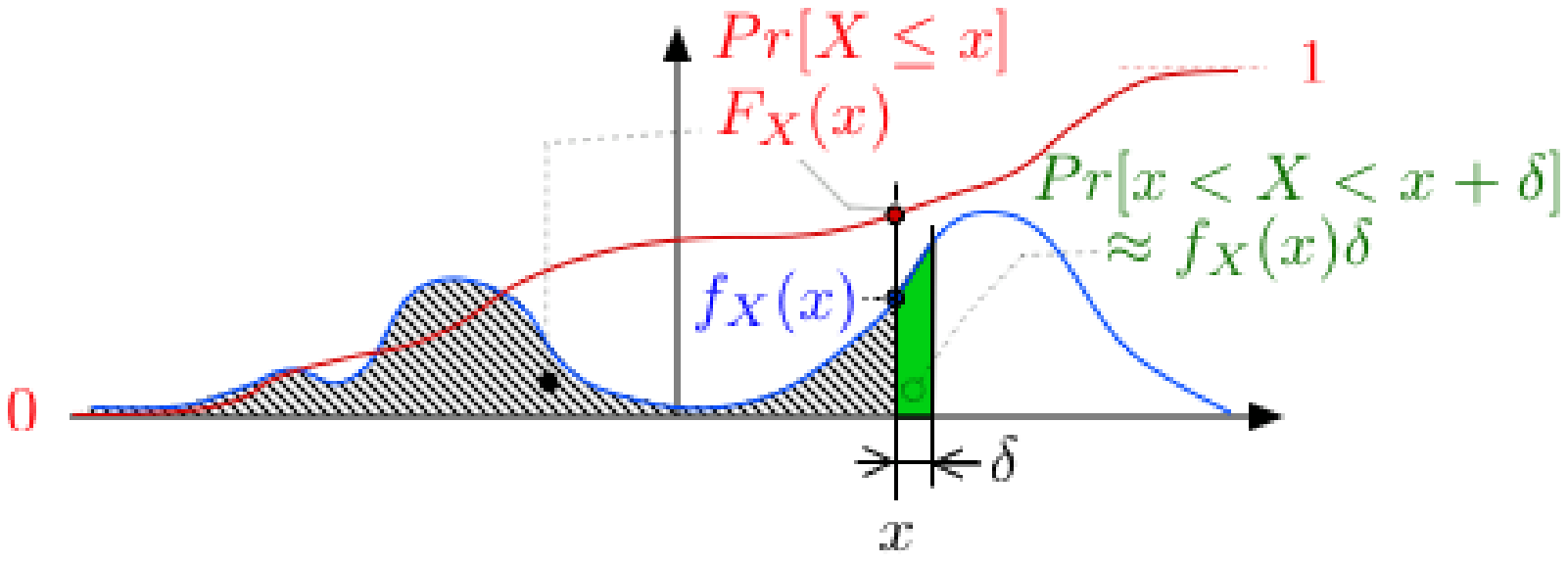


A Picture



The pdf $f_X(x)$ is a nonnegative function that integrates to 1.

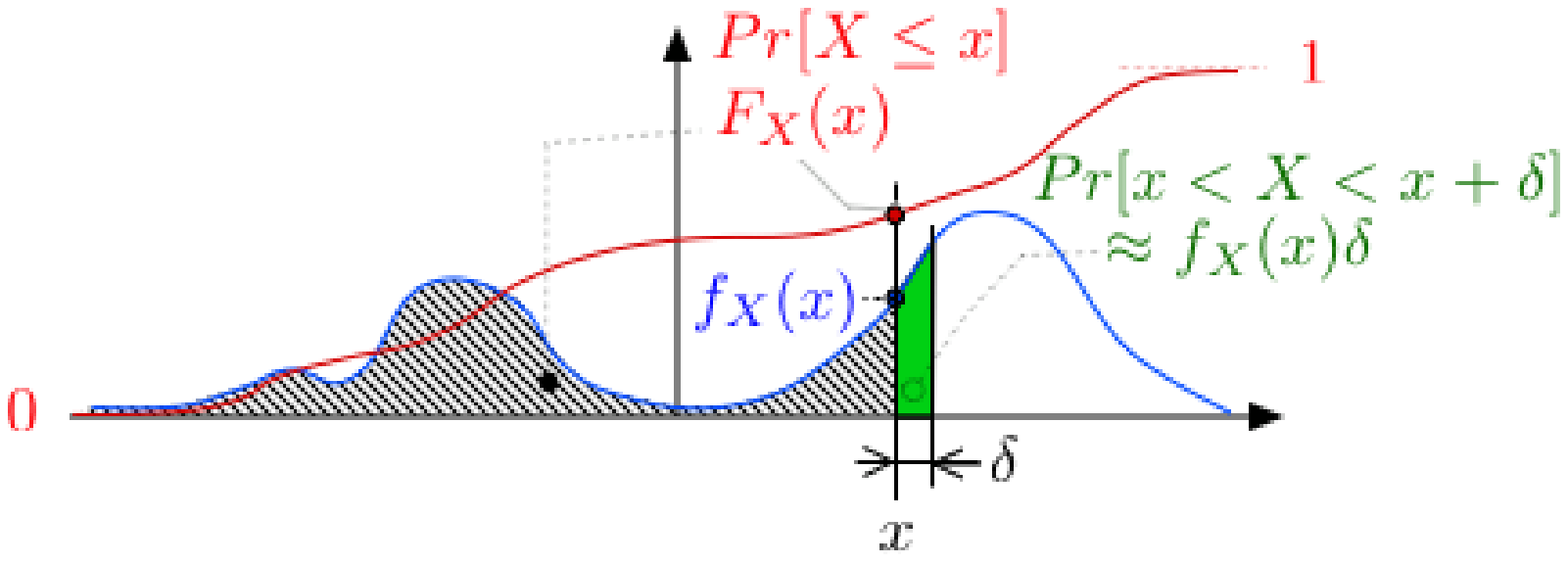
A Picture



The pdf $f_X(x)$ is a nonnegative function that integrates to 1.

The cdf $F_X(x)$ is the integral of f_X.

A Picture

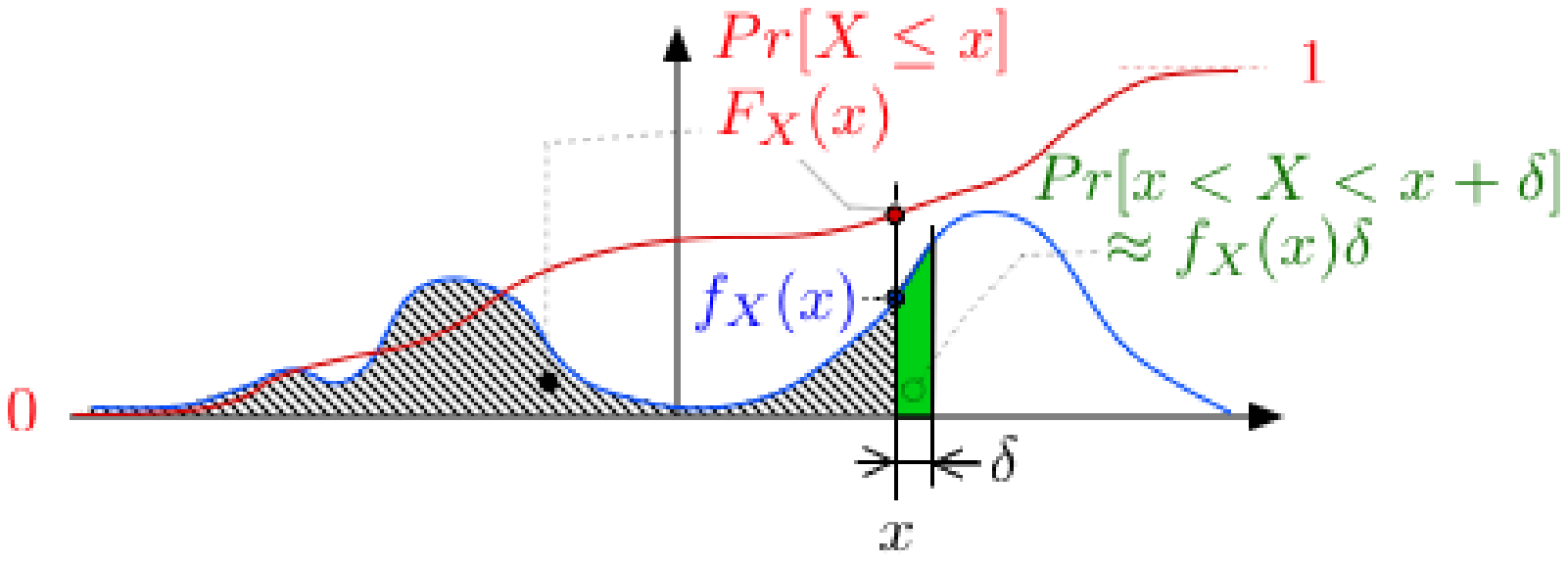


The pdf $f_X(x)$ is a nonnegative function that integrates to 1.

The cdf $F_X(x)$ is the integral of f_X.

$$Pr[x < X < x + \delta] \approx f_X(x)\delta$$

A Picture



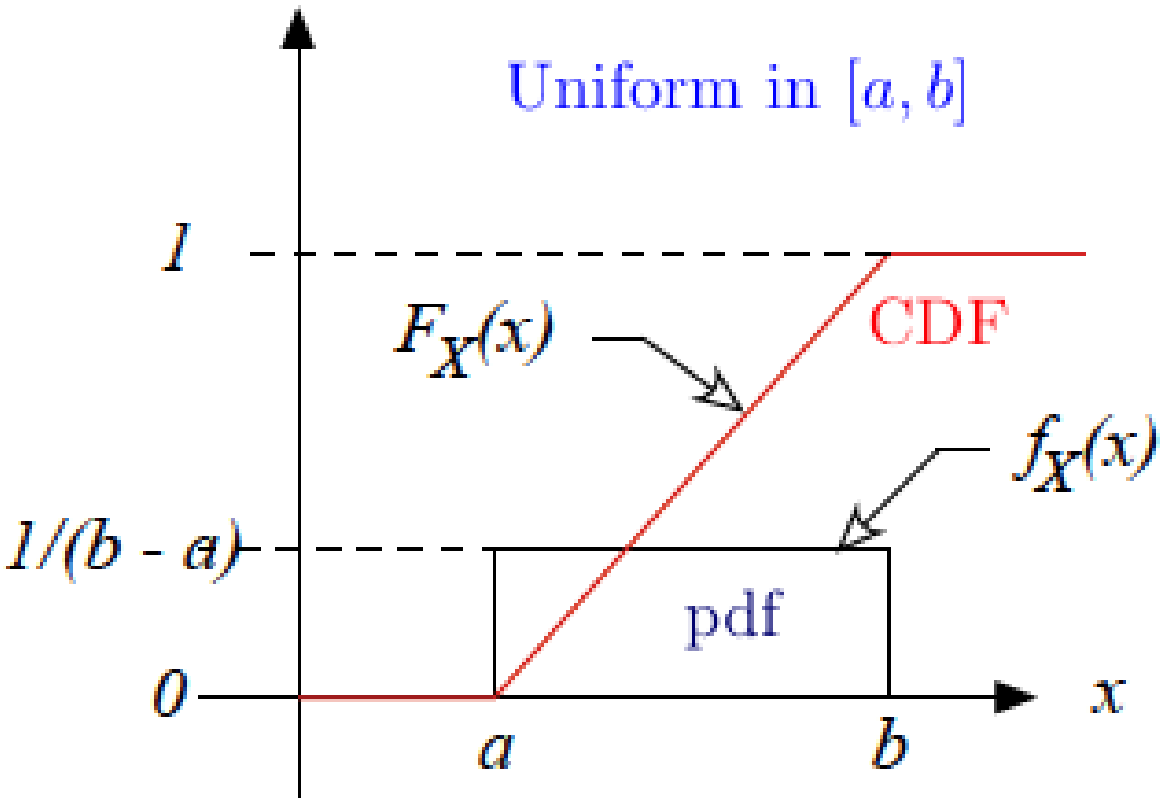
The pdf $f_X(x)$ is a nonnegative function that integrates to 1.

The cdf $F_X(x)$ is the integral of f_X.

$$Pr[x < X < x + \delta] \approx f_X(x)\delta$$

$$Pr[X \leq x] = F_x(x) = \int_{-\infty}^{x} f_X(y)dy$$

Example: $U[a, b]$



$Expo(\lambda)$

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

$Expo(\lambda)$

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

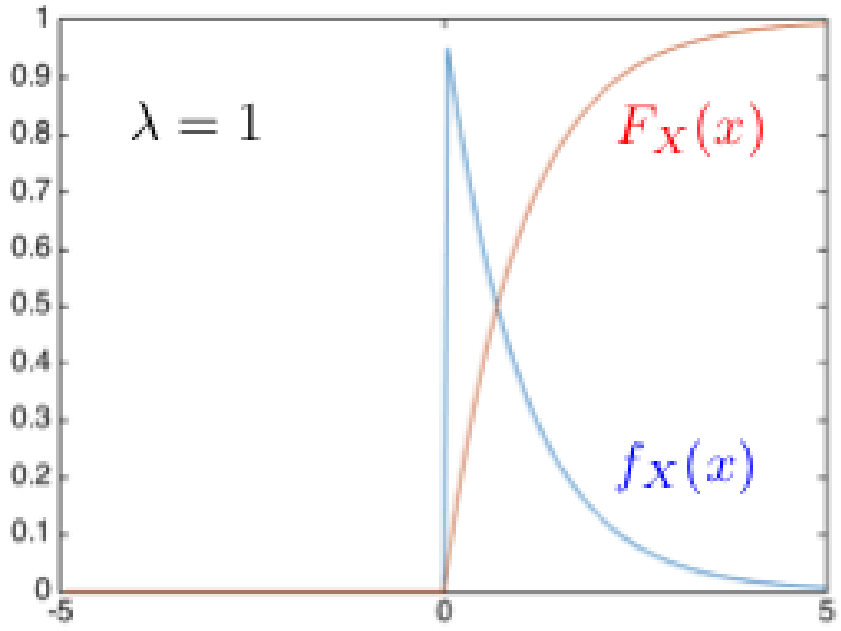
$$f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$$

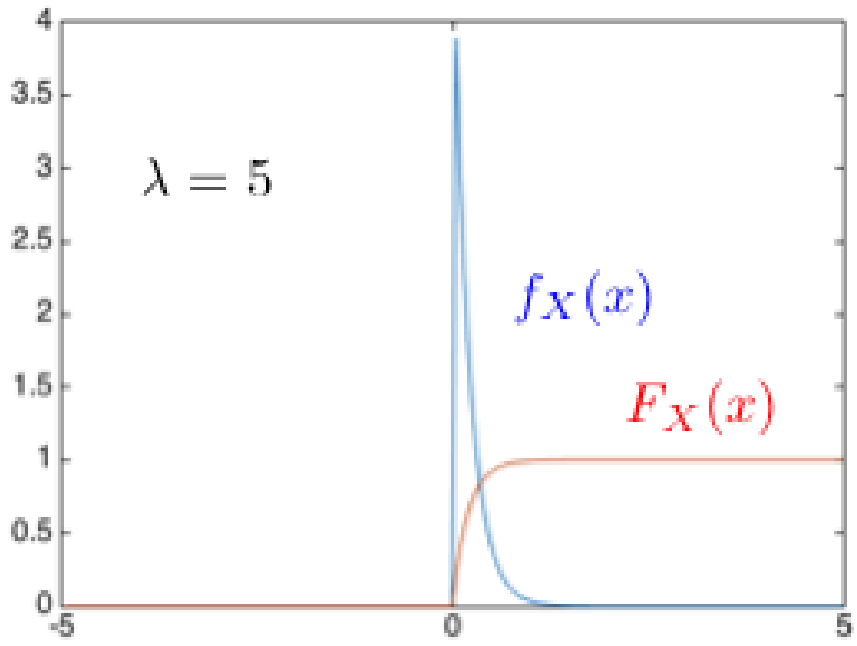
$Expo(\lambda)$

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

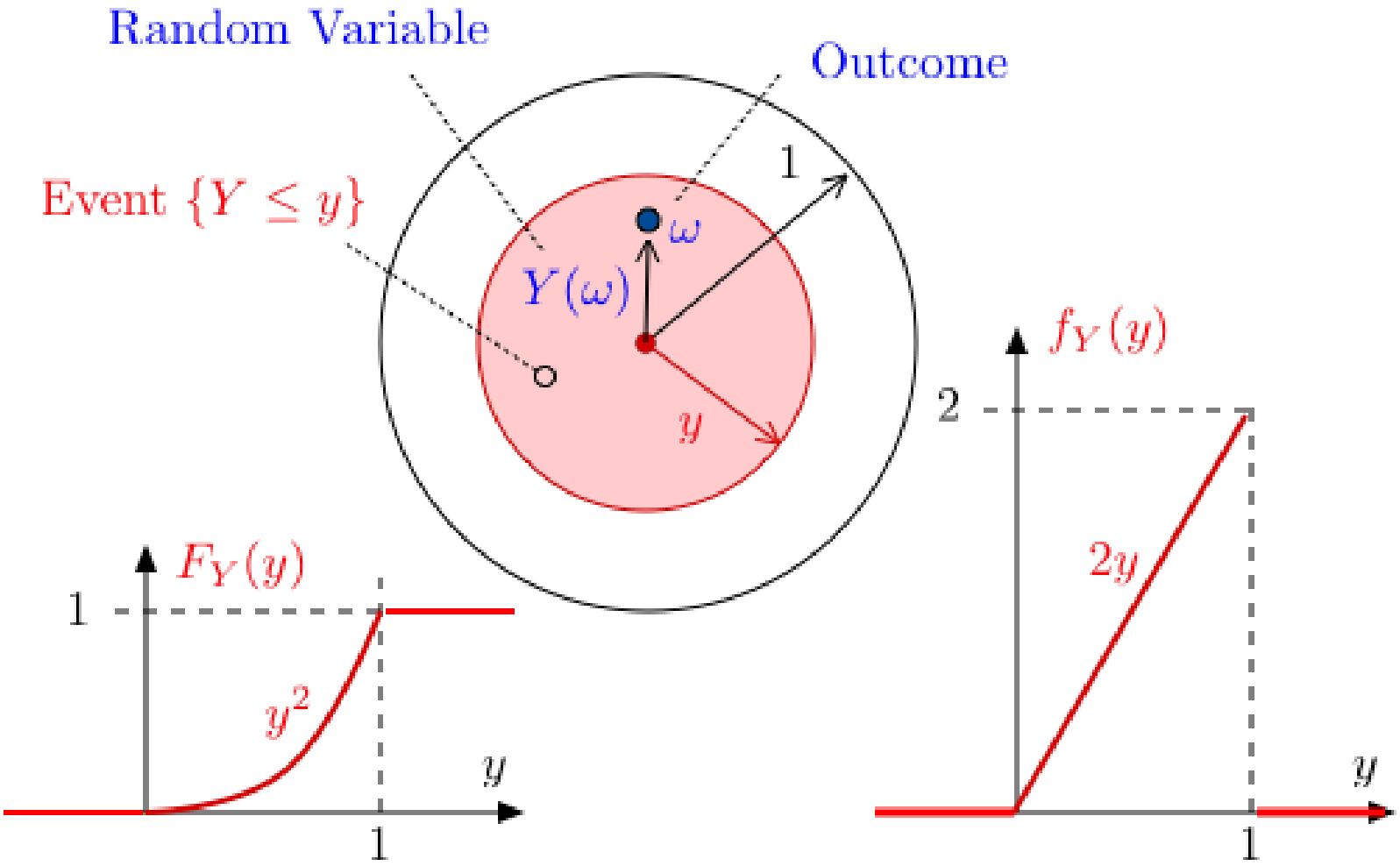
$$f_X(x) = \lambda e^{-\lambda x} 1\{x \geq 0\}$$

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





Shooting in a circle



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: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$.

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

Expectation

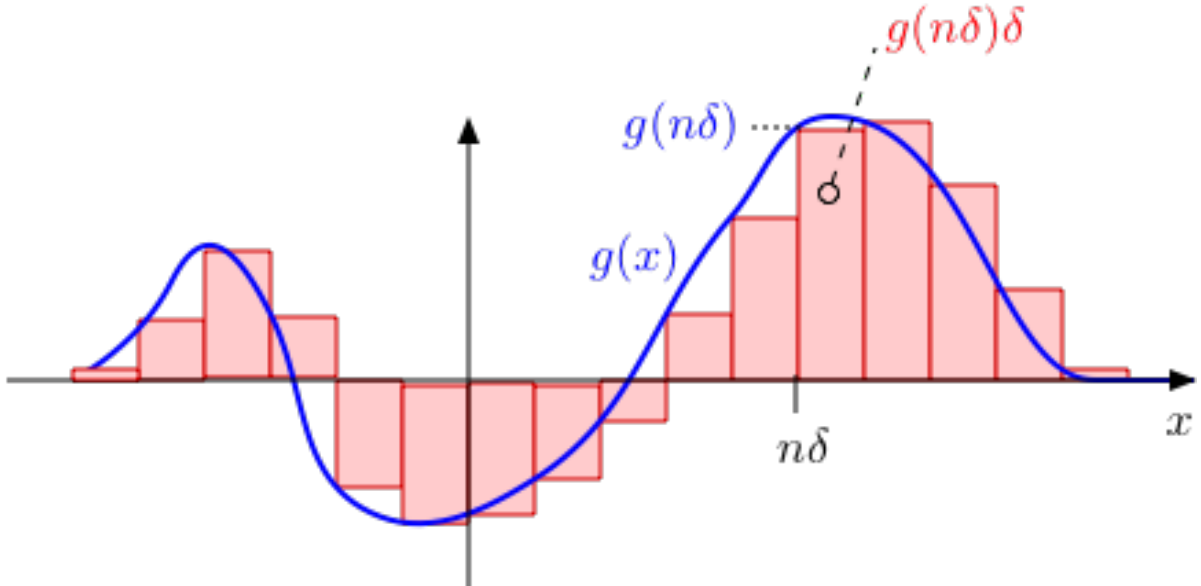
Definition The **expectation** of a random variable X with pdf $f(x)$ is defined as

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

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$. 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, $\int g(x)dx \approx \sum_n g(n\delta)\delta$ with $g(x) = x f_X(x)$.



Expectation of function of RV

Definition The expectation of a function of a random variable is defined as

Expectation of function of RV

Definition The expectation of a function of a random variable is defined as

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

Expectation of function of RV

Definition The expectation of a function of a random variable is defined as

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

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$.

Expectation of function of RV

Definition The expectation of a function of a random variable is defined as

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

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$. Then,

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

Expectation of function of RV

Definition The expectation of a function of a random variable is defined as

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

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$. Then,

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

Expectation of function of RV

Definition The expectation of a function of a random variable is defined as

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

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$. Then,

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

Indeed, $\int g(x)dx \approx \sum_n g(n\delta)\delta$ with $g(x) = h(x) f_X(x)$.

Expectation of function of RV

Definition The expectation of a function of a random variable is defined as

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

Justification: Say $X = n\delta$ w.p. $f_X(n\delta)\delta$. Then,

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

Indeed, $\int g(x)dx \approx \sum_n g(n\delta)\delta$ with $g(x) = h(x) f_X(x)$.

Fact Expectation is linear.

Variance

Definition: The **variance** of a continuous random variable X is defined as

Variance

Definition: The **variance** of a continuous random variable X is defined as

$$var[X] = E((X - E(X))^2)$$

Variance

Definition: The **variance** of a continuous random variable X is defined as

$$var[X] = E((X - E(X))^2) = E(X^2) - (E(X))^2$$

Variance

Definition: The **variance** of a continuous random variable X is defined as

$$\begin{aligned} var[X] &= E((X - E(X))^2) = E(X^2) - (E(X))^2 \\ &= \int_{-\infty}^{\infty} x^2 f(x) dx - \left(\int_{-\infty}^{\infty} x f(x) dx \right)^2 . \end{aligned}$$

Motivation for Gaussian Distribution

Key fact:

Motivation for Gaussian Distribution

Key fact: The sum of many small independent RVs has a Gaussian distribution.

Motivation for Gaussian Distribution

Key fact: The sum of many small independent RVs has a Gaussian distribution.

This is the Central Limit Theorem.

Motivation for Gaussian Distribution

Key fact: The sum of many small independent RVs has a Gaussian distribution.

This is the Central Limit Theorem. (See later.)

Motivation for Gaussian Distribution

Key fact: The sum of many small independent RVs has a Gaussian distribution.

This is the Central Limit Theorem. (See later.)

Examples: Binomial and Poisson suitably scaled.

Motivation for Gaussian Distribution

Key fact: The sum of many small independent RVs has a Gaussian distribution.

This is the Central Limit Theorem. (See later.)

Examples: Binomial and Poisson suitably scaled.

This explains why the Gaussian distribution

Motivation for Gaussian Distribution

Key fact: The sum of many small independent RVs has a Gaussian distribution.

This is the Central Limit Theorem. (See later.)

Examples: Binomial and Poisson suitably scaled.

This explains why the Gaussian distribution (the bell curve)

Motivation for Gaussian Distribution

Key fact: The sum of many small independent RVs has a Gaussian distribution.

This is the Central Limit Theorem. (See later.)

Examples: Binomial and Poisson suitably scaled.

This explains why the Gaussian distribution (the bell curve) shows up everywhere.

Normal Distribution.

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

Normal Distribution.

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

Normal Distribution.

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

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

Normal Distribution.

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

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

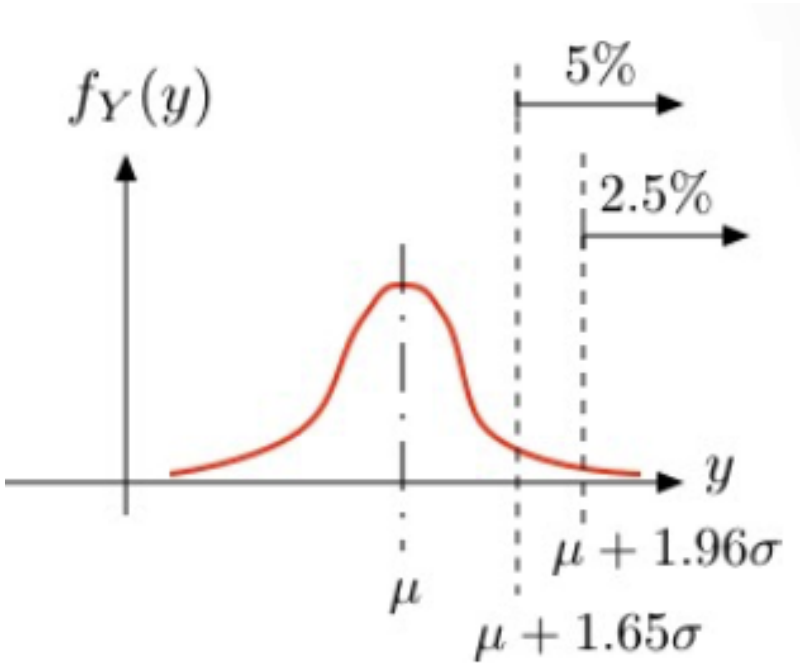
Standard normal has $\mu = 0$ and $\sigma = 1$.

Normal 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 $\sigma = 1$.

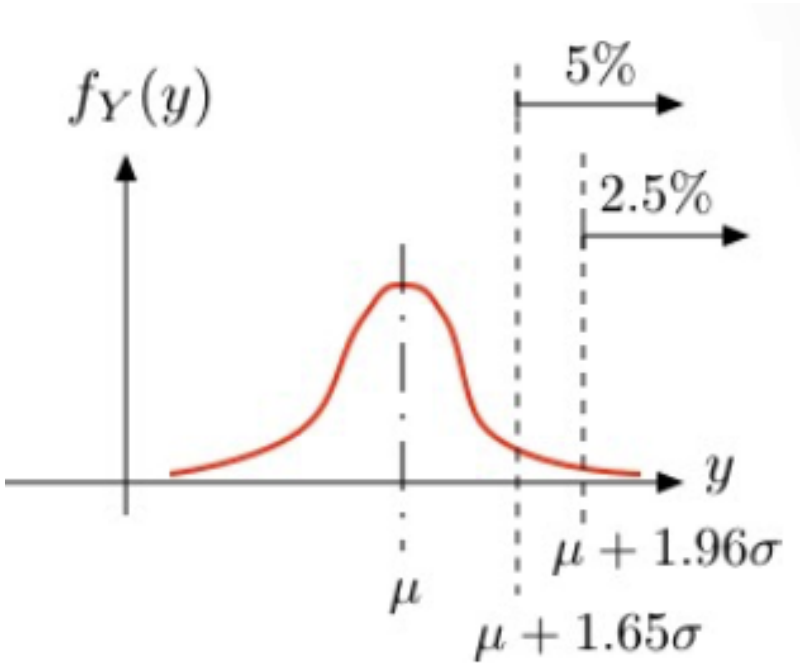


Normal Distribution.

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 $\sigma = 1$.



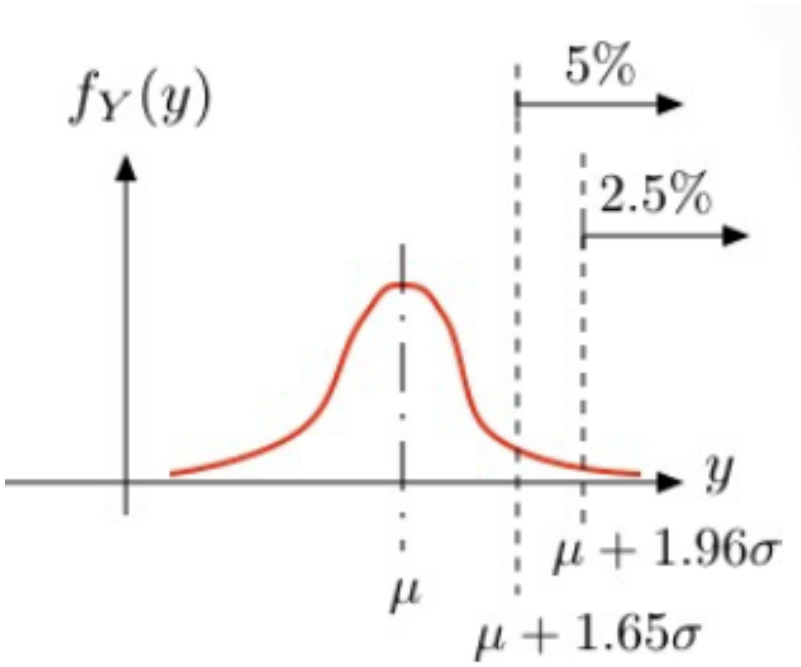
Note: $Pr[|Y - \mu| > 1.65\sigma] = 10\%$;

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

Scaling and Shifting

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

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

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$f_Y(y)dy \quad = \quad Pr[Y \in [y, y+dy]] =$$

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$f_Y(y)dy \quad = \quad Pr[Y \in [y, y+dy]] = Pr[\mu + \sigma X \in [y, y+dy]]$$

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$\begin{aligned} f_Y(y)dy &= Pr[Y \in [y, y+dy]] = Pr[\mu + \sigma X \in [y, y+dy]] \\ &= Pr[\sigma X \in [y-\mu, y-\mu+dy]] \end{aligned}$$

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$\begin{aligned} f_Y(y)dy &= Pr[Y \in [y, y+dy]] = Pr[\mu + \sigma X \in [y, y+dy]] \\ &= Pr[\sigma X \in [y-\mu, y-\mu+dy]] \\ &= Pr[X \in [\frac{y-\mu}{\sigma}, \frac{y-\mu}{\sigma} + \frac{dy}{\sigma}]] \end{aligned}$$

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$\begin{aligned} f_Y(y)dy &= Pr[Y \in [y, y+dy]] = Pr[\mu + \sigma X \in [y, y+dy]] \\ &= Pr[\sigma X \in [y-\mu, y-\mu+dy]] \\ &= Pr[X \in [\frac{y-\mu}{\sigma}, \frac{y-\mu}{\sigma} + \frac{dy}{\sigma}]] \\ &= f_X(\frac{y-\mu}{\sigma})\frac{dy}{\sigma} \end{aligned}$$

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$\begin{aligned} f_Y(y)dy &= Pr[Y \in [y, y+dy]] = Pr[\mu + \sigma X \in [y, y+dy]] \\ &= Pr[\sigma X \in [y-\mu, y-\mu+dy]] \\ &= Pr[X \in [\frac{y-\mu}{\sigma}, \frac{y-\mu}{\sigma} + \frac{dy}{\sigma}]] \\ &= f_X(\frac{y-\mu}{\sigma})\frac{dy}{\sigma} = \frac{1}{\sigma} f_X(\frac{y-\mu}{\sigma})dy \end{aligned}$$

Scaling and Shifting

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

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

Proof: $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$. Now,

$$\begin{aligned}
f_Y(y)dy &= Pr[Y \in [y, y+dy]] = Pr[\mu + \sigma X \in [y, y+dy]] \\
&= Pr[\sigma X \in [y-\mu, y-\mu+dy]] \\
&= Pr[X \in [\frac{y-\mu}{\sigma}, \frac{y-\mu}{\sigma} + \frac{dy}{\sigma}]] \\
&= f_X(\frac{y-\mu}{\sigma})\frac{dy}{\sigma} = \frac{1}{\sigma} f_X(\frac{y-\mu}{\sigma})dy \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} \exp\{-\frac{(y-\mu)^2}{2\sigma^2}\}dy. \quad \square
\end{aligned}$$

Expectation, Variance.

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

$$E[Y] = \mu$$

Expectation, Variance.

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

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

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathscr{N}(0,1)$ since $Y = \mu + \sigma X$,....

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathscr{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathscr{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathscr{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

$$var[X] = E[X^2]$$

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathscr{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

$$var[X] = E[X^2] = \int x^2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} dx$$

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathcal{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

$$\begin{aligned} var[X] &= E[X^2] = \int x^2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} dx \\ &= -\frac{1}{\sqrt{2\pi}} \int x d \exp\{-\frac{x^2}{2}\} \end{aligned}$$

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathscr{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

$$\begin{aligned} var[X] &= E[X^2] = \int x^2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} dx \\ &= -\frac{1}{\sqrt{2\pi}} \int x d \exp\{-\frac{x^2}{2}\} = \frac{1}{\sqrt{2\pi}} \int \exp\{-\frac{x^2}{2}\} dx \text{ by IBP} \end{aligned}$$

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathcal{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

$$\begin{aligned} var[X] &= E[X^2] = \int x^2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} dx \\ &= -\frac{1}{\sqrt{2\pi}} \int x d \exp\{-\frac{x^2}{2}\} = \frac{1}{\sqrt{2\pi}} \int \exp\{-\frac{x^2}{2}\} dx \text{ by IBP} \\ &= \int f_X(x) dx \end{aligned}$$

Expectation, Variance.

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

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

Proof: It suffices to show the result for $X = \mathcal{N}(0,1)$ since $Y = \mu + \sigma X$,....

Thus, $f_X(x) = \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\}$.

First note that $E[X] = 0$, by symmetry.

$$\begin{aligned} var[X] &= E[X^2] = \int x^2 \frac{1}{\sqrt{2\pi}} \exp\{-\frac{x^2}{2}\} dx \\ &= -\frac{1}{\sqrt{2\pi}} \int x d \exp\{-\frac{x^2}{2}\} = \frac{1}{\sqrt{2\pi}} \int \exp\{-\frac{x^2}{2}\} dx \text{ by IBP} \\ &= \int f_X(x) dx = 1. \quad \square \end{aligned}$$

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

$E(A'_n)$

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

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

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

$E(A'_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu) = 0.$

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

$E(A'_n) = \frac{1}{\sigma/\sqrt{n}}(E(A_n) - \mu) = 0.$

$Var(A'_n)$

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

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

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

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

Central limit theorem.

Law of Large Numbers: For any set of independent identically distributed random variables, X_i, $A_n = \frac{1}{n}\sum X_i$ "tends to the mean."

Say X_i have expecation $\mu = E(X_i)$ and variance σ^2.

Mean of A_n is μ, and variance is $\frac{\sigma^2}{n}$.

Let $A'_n = \frac{A_n - \mu}{\sigma/\sqrt{n}}$.

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

Central limit theorem: As n goes to infinity the distribution of A'_n approaches the standard normal distribution.

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

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$.

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$.

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.

CLT states that

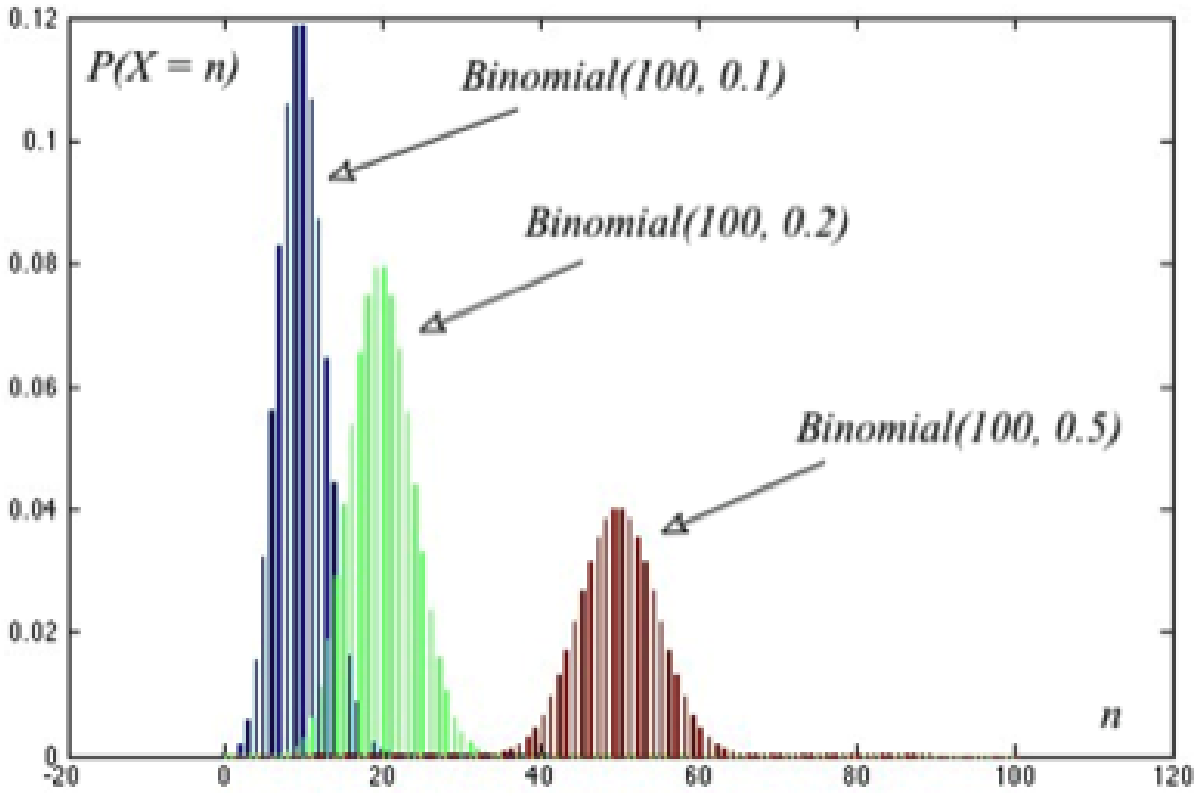
$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.

CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$

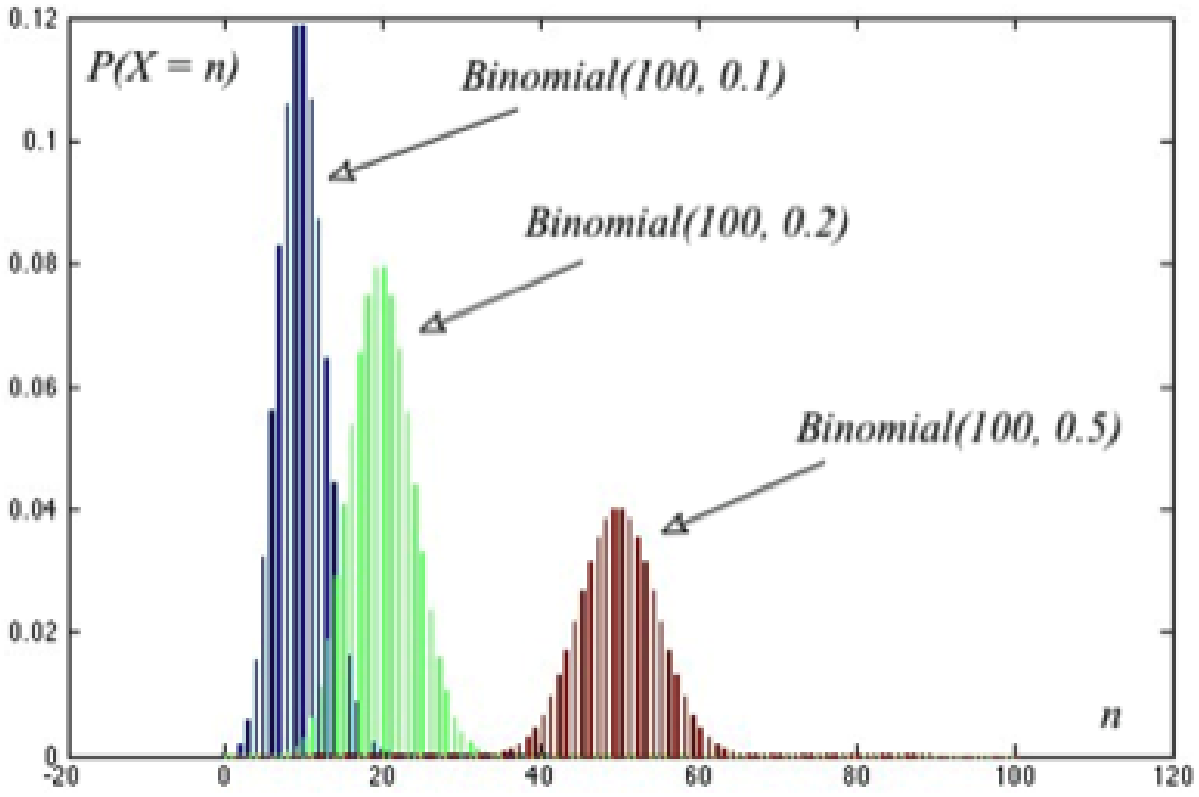


Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.

CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$



Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$.

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$.

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n, p)$.

CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.

CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$

Thus,

$$Pr[|\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}}| \geq 2] \approx 5\%.$$

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.

CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$

Thus,

$$Pr[|\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}}| \geq 2] \approx 5\%.$$

Hence,

$$Pr[|\frac{X_1 + \cdots + X_n - np}{(1/2)\sqrt{n}}| \geq 2] \leq 5\%.$$

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. Thus, $X_1 + \cdots + X_n = B(n,p)$.

CLT states that

$$\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}} \to \mathcal{N}(0,1).$$

Thus,

$$Pr[|\frac{X_1 + \cdots + X_n - np}{\sqrt{p(1-p)n}}| \geq 2] \approx 5\%.$$

Hence,

$$Pr[|\frac{X_1 + \cdots + X_n - np}{(1/2)\sqrt{n}}| \geq 2] \leq 5\%.$$

This implies that

$$Pr[p \in [\frac{X_1 + \cdots + X_n}{n} - \frac{1}{\sqrt{n}}, \frac{X_1 + \cdots + X_n}{n} + \frac{1}{\sqrt{n}}] \geq 95\%.$$

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$.

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. We just saw that

$$Pr[p \in [\frac{X_1 + \cdots + X_n}{n} - \frac{1}{\sqrt{n}}, \frac{X_1 + \cdots + X_n}{n} + \frac{1}{\sqrt{n}}] \geq 95\%.$$

Coins and normal..

Let $X_1, X_2, \ldots$ be i.i.d. $B(p)$. We just saw that

$$Pr[p \in [\frac{X_1 + \cdots + X_n}{n} - \frac{1}{\sqrt{n}}, \frac{X_1 + \cdots + X_n}{n} + \frac{1}{\sqrt{n}}] \geq 95\%.$$

Hence,

$$[\frac{X_1 + \cdots + X_n}{n} - \frac{1}{\sqrt{n}}, \frac{X_1 + \cdots + X_n}{n} + \frac{1}{\sqrt{n}}] \text{ is a } 95\% - \text{CI for } p.$$

CI for Mean

Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2.

CI for Mean

Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$

Recall that $E[A_n] = \mu$ and $var[A_n] = \frac{\sigma^2}{n}$.

CI for Mean

Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$

Recall that $E[A_n] = \mu$ and $var[A_n] = \frac{\sigma^2}{n}$. The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0,1) \text{ as } n \to \infty.$$

CI for Mean

Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$

Recall that $E[A_n] = \mu$ and $var[A_n] = \frac{\sigma^2}{n}$. The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0,1) \text{ as } n \to \infty.$$

Thus, for $n \gg 1$, one has

$$Pr[-2 \leq |\frac{A_n - \mu}{\sigma/\sqrt{n}}| \leq 2] \approx 95\%.$$

CI for Mean

Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$

Recall that $E[A_n] = \mu$ and $var[A_n] = \frac{\sigma^2}{n}$. The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0,1) \text{ as } n \to \infty.$$

Thus, for $n \gg 1$, one has

$$Pr[-2 \leq |\frac{A_n - \mu}{\sigma/\sqrt{n}}| \leq 2] \approx 95\%.$$

Equivalently,

$$Pr[\mu \in [A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]] \approx 95\%.$$

CI for Mean

Let $X_1, X_2, \ldots$ be i.i.d. with mean μ and variance σ^2. Let

$$A_n = \frac{X_1 + \cdots + X_n}{n}.$$

Recall that $E[A_n] = \mu$ and $var[A_n] = \frac{\sigma^2}{n}$. The CLT states that

$$\frac{A_n - \mu}{\sigma/\sqrt{n}} \to \mathcal{N}(0,1) \text{ as } n \to \infty.$$

Thus, for $n \gg 1$, one has

$$Pr[-2 \leq |\frac{A_n - \mu}{\sigma/\sqrt{n}}| \leq 2] \approx 95\%.$$

Equivalently,

$$Pr[\mu \in [A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}]] \approx 95\%.$$

That is,

$$[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] \text{ is a } 95\% - \text{CI for } \mu.$$

Summary

$\boxed{\text{Gaussian and CLT}}$

1. Gaussian: $\mathscr{N}(\mu, \sigma^2) : f_X(x) = \ldots$ "bell curve"
2. CLT: X_n i.i.d. $\implies \frac{A_n - \mu}{\sigma/\sqrt{n}} \to \mathscr{N}(0,1)$
3. CI: $[A_n - 2\frac{\sigma}{\sqrt{n}}, A_n + 2\frac{\sigma}{\sqrt{n}}] = 95\%$-CI for μ.