Statistics and Normal Distribution

MML 6.4-5

Marek Petrik

9/05/2023

Important Properties

Marginalization

$$\mathbb{P}[X = x] = \sum_{y \in \mathcal{T}} \mathbb{P}[X = x, Y = y]$$

Conditional probability

$$\mathbb{P}[X = x \mid Y = y] = \frac{\mathbb{P}[X = x, Y = y]}{\mathbb{P}[Y = y]}$$

Product rule

$$\mathbb{P}[X = x, Y = y] = \mathbb{P}[X = x \mid Y = y] \cdot \mathbb{P}[Y = y]$$

Independent random variables

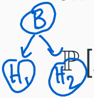
$$\mathbb{P}[X = x, Y = y] = \mathbb{P}[X = x] \cdot \mathbb{P}[Y = y]$$

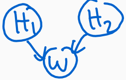
Bayes Theorem

$$\mathbb{P}[X = x \mid Y = y] = \frac{\mathbb{P}[Y = y \mid X = x] \cdot \mathbb{P}[X = x]}{\mathbb{P}[Y = y]}$$

Bayes Nets

Random variables $X, Y : \Omega \to \mathbb{R}$ are **independent conditionally** on a random variable Z if for all $x, y, z \in \mathbb{R}$:





$$\mathbb{P}[X = x, Y = y \mid Z = z] = \mathbb{P}[X = x \mid Z = z] \cdot \mathbb{P}[Y = y \mid Z = z]$$

Examples?

$\boxed{\text{Dependent and Conditionally Indep.}}$

Biased coin

$B : \Omega \to \{b, u\}$ b comes up head 70% of the time

$H_1 : \Omega \to \{h, t\}$

$H_2 : \Omega \to \{h, t\}$

$\boxed{\text{Same coin}}$

$\mathbb{P}[H_2 = h \mid H_1 = h] \neq \mathbb{P}[H_2 = h]$

$\mathbb{P}[H_1 = h, H_2 = h \mid B = b] =$

$\mathbb{P}[H_1 = h \mid B = b]\, \mathbb{P}[H_2 = h \mid B = b]$

Independent but conditionally dependent

Unbiased coin

H_1, H_2

$\mathbb{P}[H_1 = h, H_2 = h] = \mathbb{P}[H_1 = h]\, \mathbb{P}[H_2 = h]$

$W : \Omega \to \{w, \ell\}$ win $H_1 = h, H_2 = h$

$\mathbb{P}[H_1 = h, H_2 = h \mid W = w] \neq$

$\mathbb{P}[H_1 = h \mid W = w]\, \mathbb{P}[H_2 = h \mid W = w]$

Continuous Probability Distributions

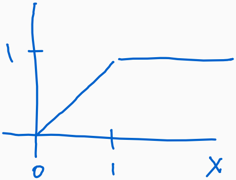
- **Normal**: common because of central limit theorem
- **Laplace**: Extreme weather events
- **Multivariate normal**: Height and weight

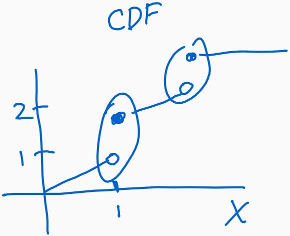
See Wikipedia for their properties

Cumulative Distribution Function

$$F_X(x) \ = \ \mathbb{P}\left[X \le x\right] = P\left(\{\omega \in \Omega \mid X(\omega) \le x\}\right)$$

Example: Uniform random variable on $[0, 1]$





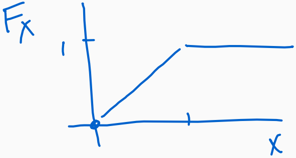
Probability Density Function PDF

Function $f_X : \mathbb{R} \to \mathbb{R}$ is a pdf of X if $\forall x$

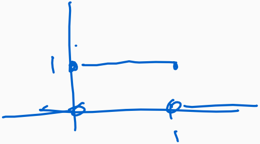
$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

$$f_X(x) = \frac{d}{dx} F_X(x) \quad \text{if } f_X \text{ is continuous}$$

CDF



PDF



$\mathbb{P}[X = 0.5] = 0$

$\mathbb{P}[X = 0.1] = 0$

Conditional Probability Density Function

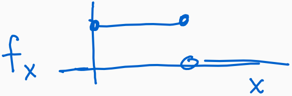
For random variables X, Y with a joint density $f_{X,Y} : \mathbb{R}^2 \to \mathbb{R}$, the conditional density is (when $f_X(x) > 0$):

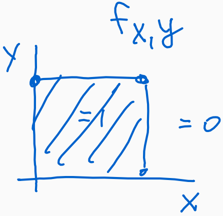
$$f_{Y|X}(y \mid x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

$\mathbb{P}[X = x \mid Y = y]$
undefined

Y is continuous
$\mathbb{P}[Y = y] = 0$

$Y|X \equiv Y$ conditional on X





Today

1. Expected value
2. Variance, covariance, and correlation
3. Normal distribution

Expected Value (Mean)

Random variable X (discrete and continuous, see Lebesgue integrals)

$$\mathbb{E}[X] = \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}[X = x], \qquad \mathbb{E}[X] = \int_\Omega X dP$$

$H(J) = 1$
$H(M) = 1$
$H(E) = 2$

$P(\{J\}) = \frac{1}{2}$
$P(\{M\}) = P(\{E\}) = \frac{1}{4}$

$$\mathbb{E}[H] = 1 \cdot \left(\frac{1}{2} + \frac{1}{4}\right) + 2\left(\frac{1}{4}\right) =$$

$$= 1 \cdot \frac{3}{4} + 2 \cdot \frac{1}{4} = \frac{5}{4}$$

Expected Value of Function

Random variable X (discrete and continuous) and a function g

$$\mathbb{E}\left[\underbrace{g(X)}_{Y=g(X)}\right] = \sum_{g'\in\mathcal{G}} g' \cdot \mathbb{P}\left[g(X) = g'\right], \qquad \mathbb{E}\left[X\right] = \int_{\Omega} g(X)dP$$

Law of the Unconscious Statistician proves that:

$$\mathbb{E}[g(X)] = \sum_{x\in\mathcal{X}} g(x)\,\mathbb{P}[X=x]$$

$$\mathbb{E}[H^2]$$

Expected Value: Properties

Assume real-valued random variables X and Y

$$\mathbb{E}[c \cdot X] = c \cdot \mathbb{E}[X]$$

$\mathbb{E}[2 \cdot H] = 2\,\mathbb{E}[H]$

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y]$$

$\mathbb{E}$ is a linear <u>operator</u>

When X and Y are **independent**:

$$\mathbb{E}[X \cdot Y] = \mathbb{E}[X] \cdot \mathbb{E}[Y]$$

$\mathbb{E}[X \cdot Y] =$

$= \sum_{x,y} x \cdot y \; p(x,y) = \sum_{x,y} x \cdot y \; p(x)\, p(y) = \left(\sum_x x\, p(x)\right)\left(\sum_y y\, p(y)\right) = \mathbb{E}[X] \cdot \mathbb{E}[Y]$

Conditional Expectation

https://en.wikipedia.org/wiki/Conditional_expectation

Conditioning on an event $A \in \mathcal{F}$ (discrete r.v.)

$$\mathbb{E}[X \mid A] \quad = \quad \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}[X = x \mid A]$$

$\mathbb{E}[H \mid W \geq 150]$

Conditioning on a random variable Y: $\mathbb{E}[H \mid W]$ is a r.v.

$$\mathbb{E}[X \mid Y] : \Omega \to \mathbb{R}$$

defined as (discrete)

$$\mathbb{E}[X \mid Y](\omega) \quad = \quad \sum_{x \in \mathcal{X}} x \cdot \mathbb{P}[X = x \mid Y = Y(\omega)]$$

$\mathbb{E}[X] = \mathbb{E}[\,\mathbb{E}[X \mid Y]\,]$

Expected Utility Theory

Two possible investments with profits X_1 and X_2

X_1

Choose investment one if

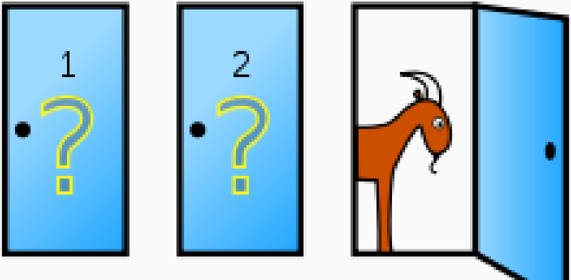
$$\underline{\mathbb{E}[X_1]} \geq \mathbb{E}[X_2]$$

Utility function: $u : \mathbb{R} \to \mathbb{R}$

Pick X_1 if $\mathbb{E}[u(X_1)] \geq \mathbb{E}[u(X_2)]$

Monty Hall Problem

https://en.wikipedia.org/wiki/Monty_Hall_problem



1. Pick a door X

Stick with the door : win $\frac{1}{3}$

Switch : win $\frac{2}{3}$

Two Envelopes Paradox

https://en.wikipedia.org/wiki/Two_envelopes_problem

I have two envelopes, one has X dollars, another has $2 \cdot X$ dollars

After choosing one, would you want to switch?

z in envelope

$$\begin{array}{ll} & p \\ 2z & \frac{1}{2} \\ \frac{1}{2}z & \frac{1}{2} \end{array}$$

$$\mathbb{E}[X_2] = 2z\frac{1}{2} + \frac{1}{2}z\frac{1}{2} = z + \frac{1}{4}z = \frac{5}{4}z$$

y in envelope

$$\mathbb{E}[X_3] = \frac{5}{4}y$$

Variance and Standard Deviation

Variance : 2nd central moment

$$\mathbb{V}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$$

Standard deviation

$$\text{sd}[X] = \sqrt{\mathbb{V}[X]} = \sqrt{\mathbb{E}\left[X - \mathbb{E}[X]^2\right]}$$

Why standard deviation?

H = in meters

$\mathbb{V}[H] = m^2$

$\text{sd}[H] = m$

Variance: Another Representation

$$\mathbb{V}[X] = \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

$$\mathbb{V}[X] = \mathbb{E}[(X - \mathbb{E}[X])^2] = \mathbb{E}[X^2 - 2X\,\mathbb{E}[X] + \mathbb{E}[X]^2] =$$

$$= \mathbb{E}[X^2] - 2\,\mathbb{E}[X]\,\mathbb{E}[X] + \mathbb{E}[X]^2$$

$$= \mathbb{E}[X^2] - \mathbb{E}[X]^2$$

Covariance

$$\mathrm{Cov}[X, Y] = \mathbb{E}[\underline{(X - \mathbb{E}[X])} \cdot \underline{(Y - \mathbb{E}[Y])}]$$

$$\mathrm{Cov}[X, Y] = \mathbb{E}[XY - Y\mathbb{E}[X] - X\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y]]$$

$$= \mathbb{E}[XY] - 2\mathbb{E}[X]\mathbb{E}[Y] + \mathbb{E}[X]\mathbb{E}[Y]$$

$$= \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y]$$

Variance and Covariance Properties

$$\text{Cov}[X, Y] = 0 \quad \text{when} \quad X, Y \text{ are independent}$$

$\text{Cov}[X,Y] \not\Rightarrow$ Independence

$$\text{Cov}[a \cdot X, Y] = \text{Cov}[X, a \cdot Y] = a \cdot \text{Cov}[X, Y]$$

$$\mathbb{V}[a \cdot X] = a^2 \cdot \mathbb{V}[X]$$

$$\mathbb{V}\left[\sum_{i \in \mathcal{I}} X_i\right] = \sum_{i \in \mathcal{I}} \sum_{j \in \mathcal{I}} \text{Cov}[X_i, X_j]$$

Variance of Independent Variables

Suppose that X and Y are independent random variables. Compute

$$\mathbb{V}[X + Y] =$$

$\mathrm{Cov}(X, X) + 2\mathrm{Cov}(X, Y) + \mathrm{Cov}(Y, Y) =$

$\mathbb{V}[X] + \mathbb{V}[Y]$

Covariance and Correlation

Covariance

$$\text{Cov}[X, Y] = \mathbb{E}[(X - \mathbb{E}[X]) \cdot (Y - \mathbb{E}[Y])]$$

Between $(-\infty, \infty)$

Correlation

corr[X,Y]

$$\cancel{\text{corr}[X]} = \frac{\text{Cov}[X, Y]}{\text{sd}[X] \cdot \text{sd}[\underset{Y}{\cancel{X}}]}$$

Between $[-1, +1]$

Correlation Coefficient

Correlation coefficient $\operatorname{corr}[X, Y]$ is between $[-1, 1]$

- 0**:** Variables are not related
- 1**:** Variables are perfectly related (same)
- -1**:** Variables are negatively related (different)

Covariance and Correlation: Properties

$$\mathrm{Cov}[X, X] = \mathbb{V}[X]$$

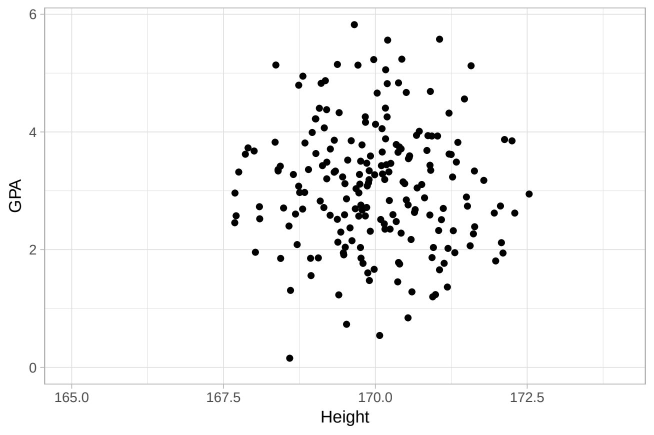
$$\mathrm{corr}[X, X] = 1$$

$sd[X] > 0$

$$\mathrm{corr}[X, Y] = 0 \text{ when } X, Y \text{ are independent}$$

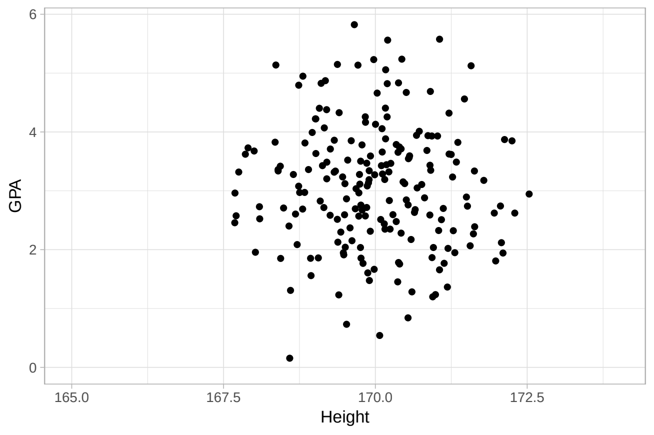
Correlation Example

Plotting code in `slides/figs/class2/plots.R`



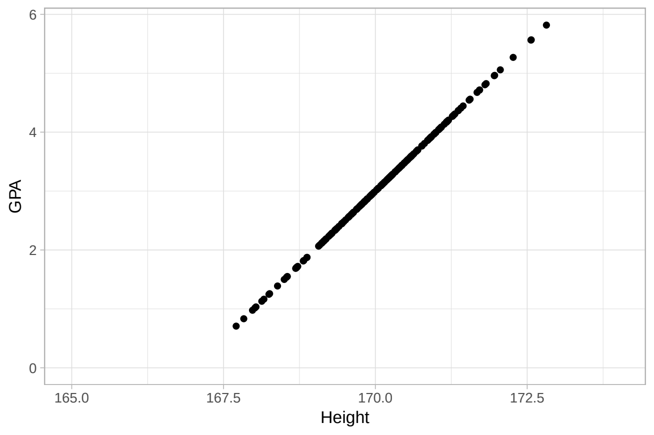
Correlation Example

Plotting code in `slides/figs/class2/plots.R`

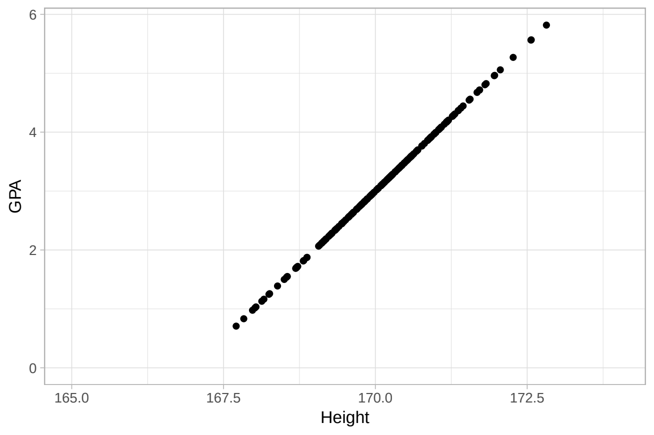


Correlation: 0

Correlation Example

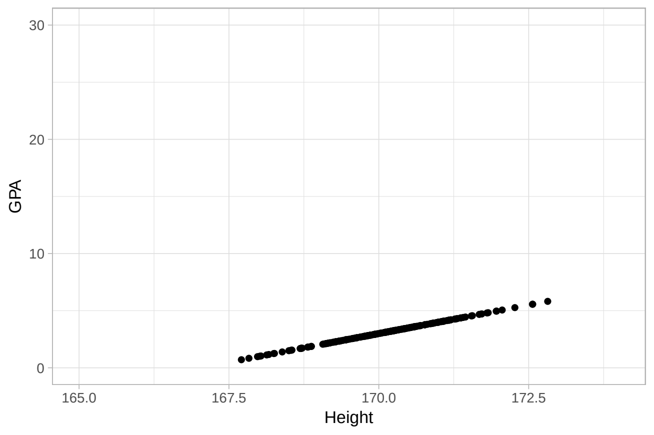


Correlation Example

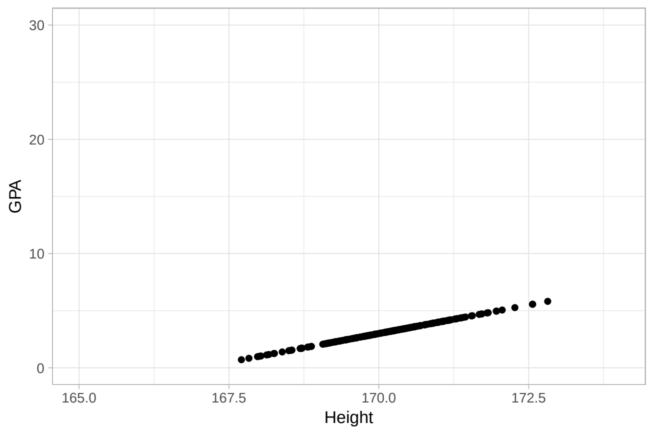


Correlation: 1

Correlation Example

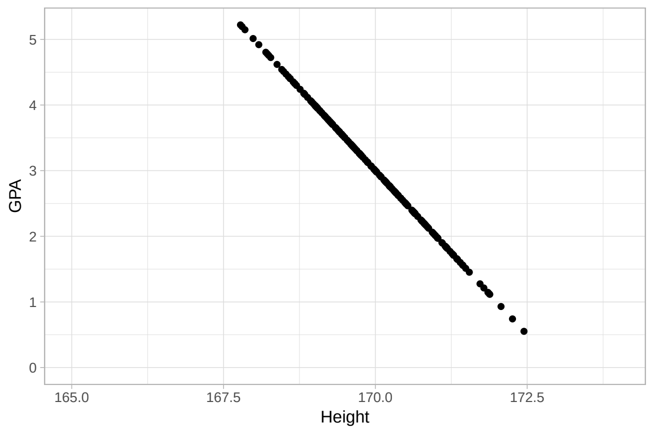


Correlation Example

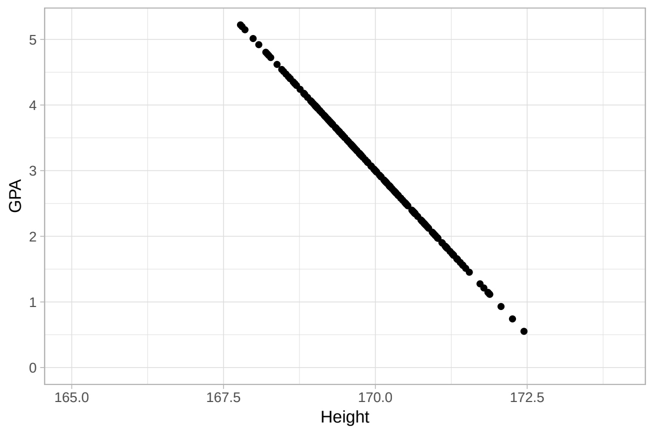


Correlation: 1

Correlation Example

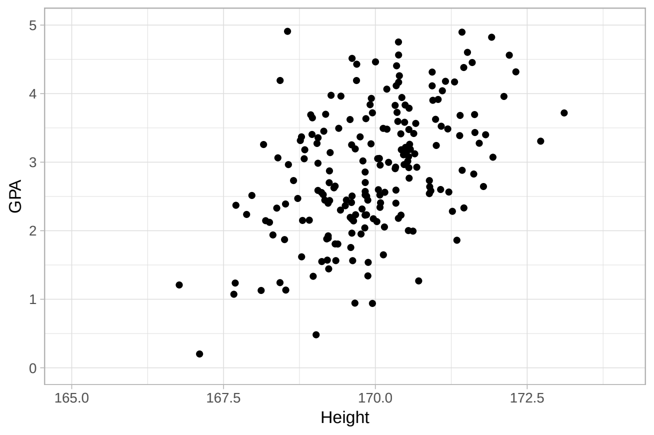


Correlation Example

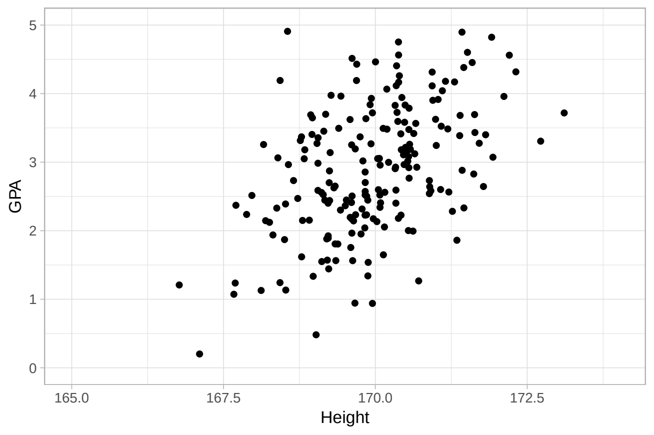


Correlation: -1

Correlation Example

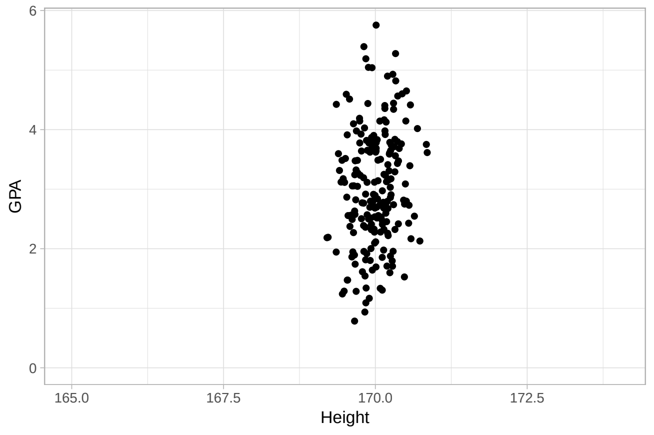


Correlation Example

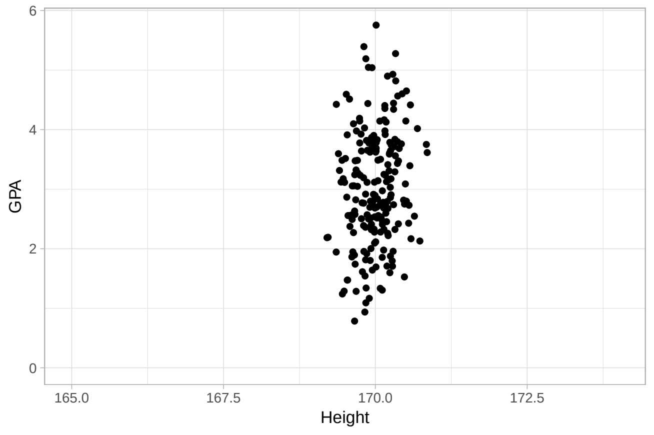


Correlation: 0.5

Correlation Example



Correlation Example

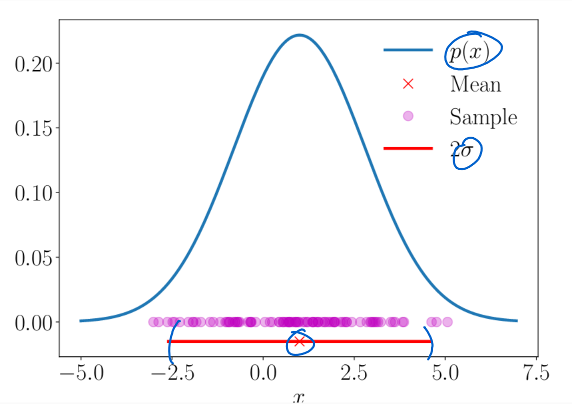


Correlation: 0.0

Univariate Normal Distribution

Density for $X \sim \mathcal{N}(\mu, \sigma^2)$

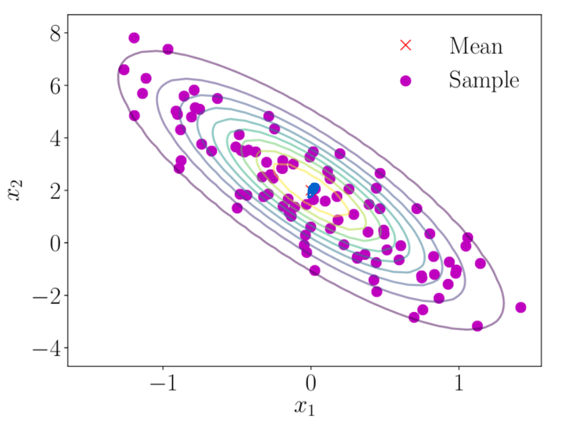
PDF
$$p(x \mid \mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{\sigma^2}\right)$$



Multivariate Normal Distribution

Joint probability density function

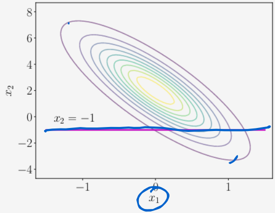
$$p(x_1, \overset{x_2}{\cancel{y}} \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \dots$$



Multivariate Normal Distribution

Joint probability density function

$$p(x, y \mid \boldsymbol{\mu}, \boldsymbol{\Sigma}) = \dots$$



(a) Bivariate Gaussian.

Marginal

Conditional