ENGG 2430 / ESTR 2004: Probability and Statistics

Spring 2019

10. Bayesian Statistics

Andrej Bogdanov

The Central Dogma of Statistics

data = independent samples from some random variable (or several random variables)

...but we don't know PDF/PMF

$\text{Poisson}(\lambda)$

$\text{Normal}(\mu, \sigma)$

Alice	**PASS**
Bob	**PASS**
Charlie	**FAIL**

$\text{Binomial}(200, p)$

Please pass me the

SALT 20%
BALL 30%
BAZOOKA 2%

?

OBAMA 1%
KIM 30%
XI 15%
MERKEL 10%

parameters λ, μ, σ, p etc. are

LAPLACE

BAYESIAN

RANDOM VAR.

PEARSON

UNKNOWN

Bayesian inference

1. Assign prior probabilities to params

2. Observe data

3. Update probabilities via Bayes' rule

Bayes' rule

$$f_{\Theta \mid X}(\theta \mid x) = \frac{f_{X \mid \Theta}(x \mid \theta)\, f_{\Theta}(\theta)}{f_X(x)}$$

POSTERIOR

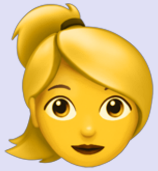
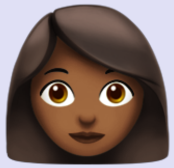
PRIOR

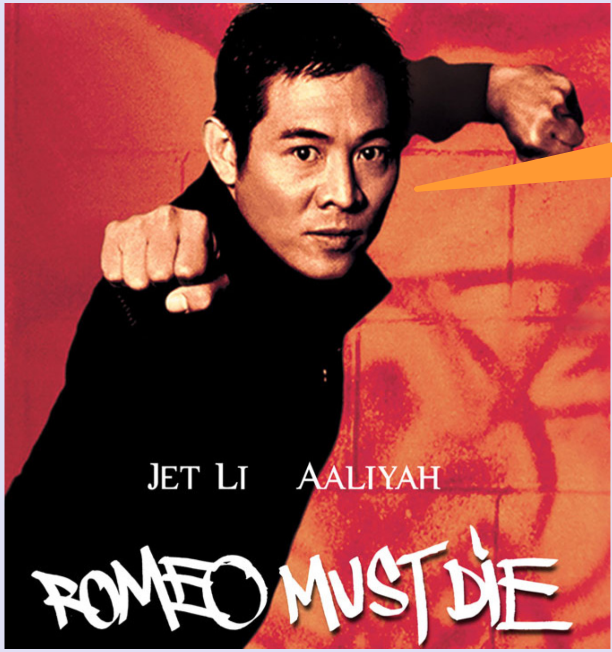
$$\propto f_{X \mid \Theta}(x \mid \theta)\, f_{\Theta}(\theta)$$

$$\boxed{f_{\Theta \mid X_1 \ldots X_n}(\theta \mid x_1 \ldots x_n) \propto f_{X_1 \mid \Theta}(x_1 \mid \theta) \ldots f_{X_n \mid \Theta}(x_n \mid \theta)\, f_{\Theta}(\theta)}$$

if $X_1, \ldots, X_n$ **are independent**

Romeo is waiting for Juliet on their first date.

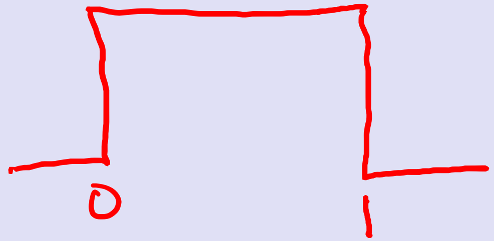
			
$X =$	$\text{Uniform}(0, .3)$	$\text{Uniform}(0, .8)$	$\text{Uniform}(0, .6)$



girls are $\text{Uniform}(0, \Theta)$ late

Romeo's model $X = \text{Uniform}(0, \Theta)$

$\Theta = \text{Uniform}(0, 1)$



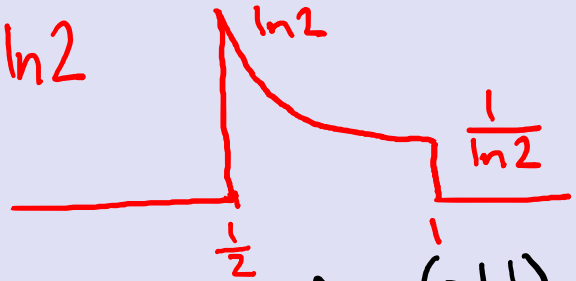
PRIOR $f_\Theta(\theta)$

On her first date, Juliet arrives ½ hour late.

$$f_{\Theta|X}\left(\theta \,\middle|\, x=\tfrac{1}{2}\right) \propto f_{X|\Theta}\left(\tfrac{1}{2} \,\middle|\, \theta\right) f_\Theta(\theta)$$

$$= \frac{1}{\theta} \cdot 1 \quad \text{IF} \quad \geq \frac{1}{2}$$

$$f_{\Theta|X}\left(\theta \,\middle|\, x=\tfrac{1}{2}\right) = \frac{\frac{1}{\theta}}{\int_{1/2}^{1} \frac{1}{\theta'}\, d\theta'} = \frac{\frac{1}{\theta}}{-\ln\frac{1}{2}} = \frac{1}{\theta \ln 2}$$



POSTERIOR $f_{\Theta|X}\left(\theta \,\middle|\, \tfrac{1}{2}\right)$

On her first 3 dates, Juliet is late by x_1, x_2, x_3 hours.

$$f_{\Theta|X_1X_2X_3}(\theta \mid x_1, x_2, x_3) \propto f_{X_1X_2X_3|\Theta}(x_1\, x_2\, x_3 \mid \theta)\, f_\Theta(\theta)$$

$$= f_{X_1|\Theta}(x_1|\theta)\, f_{X_2|\Theta}(x_2|\theta)\, f_{X_3|\Theta}(x_3|\theta)\, f_\Theta(\theta)$$

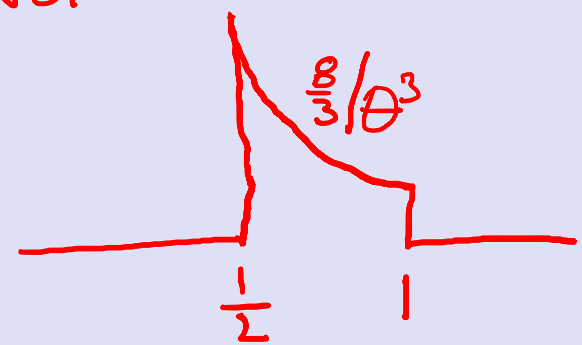
$$= \frac{1}{\theta} \cdot \frac{1}{\theta} \cdot \frac{1}{\theta} \cdot 1$$

$$x_1 \le \theta \qquad x_2 \le \theta \qquad x_3 \le \theta$$

$$= \begin{cases} \frac{1}{\theta^3} & \text{IF } x_1, x_2, x_3 \le \theta \\ 0 & \text{IF NOT} \end{cases}$$

EX. $x_1 = \frac{1}{2},\ x_2 = x_3 = \frac{1}{4}$

$$\int_{1/2}^{1} \frac{c}{\theta'^3}\, d\theta' = 1 \longrightarrow c = \frac{8}{3}$$



POSTERIOR $f_{\Theta|X}(\theta \mid \frac{1}{2}, \frac{1}{4}, \frac{1}{4})$

Three independent $\mathrm{Normal}(\Theta, 1)$ RVs take values 3.97, 4.09, 3.11. What is Θ?

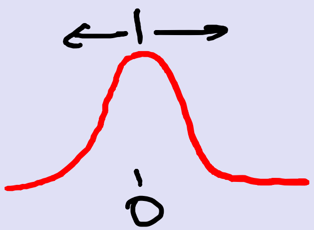
PRIOR $\Theta = \mathrm{Normal}(0,1)$

$$f_{\Theta|X_1X_2X_3}(\theta \mid x_1 x_2 x_3) \propto f(x_1|\theta) f(x_2|\theta) f(x_3|\theta) f(\theta)$$

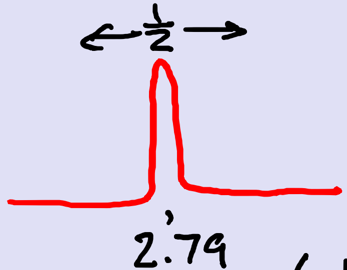
$$\propto e^{-(x_1-\theta)^2/2} e^{-(x_2-\theta)^2/2} \cdot e^{-(x_3-\theta)^2/2} \cdot e^{-\theta^2/2}$$

$$\propto e^{-\left(\frac{x_1+x_2+x_3+0}{4}-\theta\right)^2 / 2\cdot\left(\frac{1}{\sqrt{4}}\right)^2}$$

$$= \text{PDF OF Normal}\left(\frac{0+x_1+x_2+x_3}{4}, \sigma = \frac{1}{\sqrt{4}}\right)$$



PRIOR f_Θ



POSTERIOR $f_{\Theta|X_1X_2X_3}(\theta \mid 3.97, 4.09, 3.11)$

Inference for normals

$X_i = \mathrm{Normal}(\Theta, \sigma_i)$ **independent given** Θ

Θ **is** $\mathrm{Normal}(x_0, \sigma_0)$

$(\Theta \mid X_1 = x_1, \ldots, X_n = x_n)$ **is** $\mathrm{Normal}(x, \sigma)$ **where**

$$1/\sigma^2 = 1/{\sigma_0}^2 + \ldots + 1/{\sigma_n}^2$$

$$x/\sigma^2 = \frac{x_0/{\sigma_0}^2 + \ldots + x_n/{\sigma_n}^2}{n+1}$$

A coin of unknown bias flips HHTH. What is the bias?

PRIOR: BIAS θ IS Uniform$(0,1)$

$$f_{\theta|\underline{X}}(\theta \mid HHTH) \propto f_{\underline{X}|\theta}(HHTH \mid \theta) \cdot f_{\theta}(\theta)$$

$$= \theta^3 (1-\theta) \cdot 1$$

The $\mathrm{Beta}(\alpha, \beta)$ random variable

$$f_\Theta(\theta) = \frac{1}{B(\alpha, \beta)} \theta^{\alpha-1}(1-\theta)^{\beta-1} \quad \textbf{when } 0 < \theta < 1$$

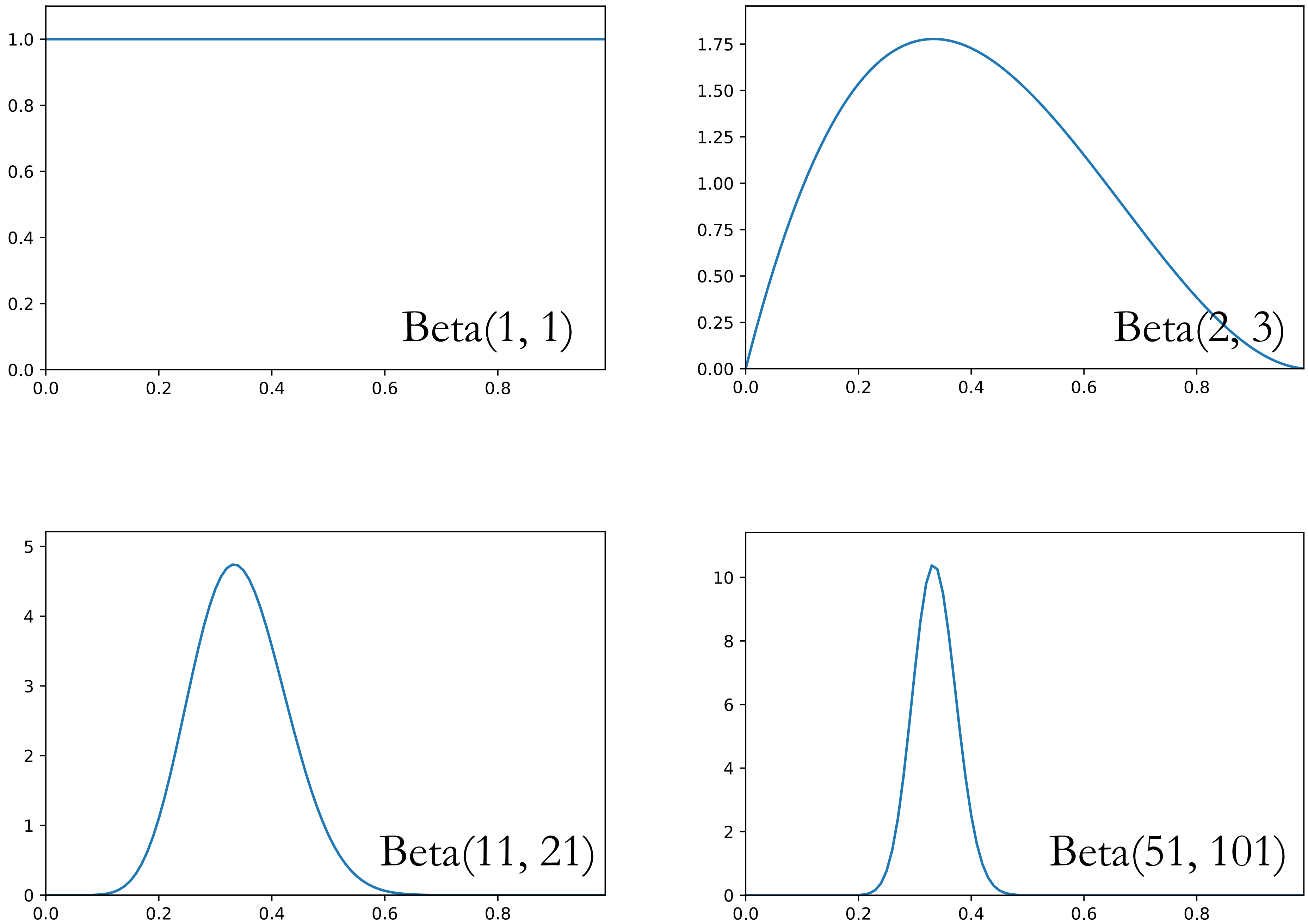
$$B(\alpha, \beta) = (\alpha - 1)!\,(\beta - 1)!\,/\,(\alpha + \beta - 1)!$$

The $\text{Beta}(\alpha, \beta)$ random variable

$$\text{Beta}(1, 1) = \text{Uniform}(0, 1)$$

Θ **is** $\text{Beta}(1, 1)$

(Θ | h **heads**, t **tails**) **is** $\text{Beta}(1 + h, 1 + t)$


Beta(1, 1)
Beta(2, 3)
Beta(11, 21)
Beta(51, 101)

CONGRATULATIONS!! 3/02

from: F48E5F6BRT@vega.ocn.ne.jp

to: andrejb@cse.cuhk.edu.hk

Dear Customer,

My name is Sandra Davis, Board of Directors of United Nations and Chief Executive Officer, effective April 16, 2018,

The United Nations {UN} has giving you extra three working days to receive your fund from Citibank Plc, New york or you will lose the opportunity for ever. So you are advised to comply immediately to avoid the cancellation of your fund, follow the instruction immediately for your own good and future

The Citibank controlling department controlling of the security transfer CODE which is (CI201), the Authentication section code of this bank concludes the verification of your file. After going through all the documents of claim received by this department with justification and verification from the global strategy United States we are completely satisfied and you have been confirmed.

The Citibank concerning wire transfers of your fund. Your letter has been referred to the (JMCB) Legal Division for Funds (US$2.8 Million Dollars) Transferred code (). We are satisfied using Electronic Wire Transfer or Swift Wire Transfer and the rights and liabilities of using of electronic and Swift fund transfer systems are defined by the Electronic Fund Transfer Act... The regulation, however, which implements this statute, Regulation E. specifically states that its provisions are inapplicable to a situation such we must ensure your Funds Transferred to your destination Bank Account between 72 hours.

Considering the volume of your payment, it is right for us to seek for the approval of some money regulatory Boards here in United States before we can carry out the Transfer of an amount of such magnitude to anybody, otherwise any such transfer will be stopped by the Authorities, and the International Monetary Fund (IMF), since your Transfer is Electronic Transfer or Swift Wire transfer is almost activated with our bank and the only thing holding the final activation of your Account are some certain Approval Documents from the concerned Authorities here in United States

NB: THIS TRANSACTION IS BEING MONITORED BY THE UNITED STATES GOVERNMENT IN ORDER TO GUARDS US FROM INTERNET IMPOSTORS.

Provide your designated bank account details for Electronic Transfer, to avoid mistake(s).

Bank Name and Address
Account Number:
Account Name:
Routing Number:
your home adress and phone number,
place of work and address.

send it the citibank remittance manager. on her email
: ombes2@gmx.com

UN gives you only 3 working days to receive your fund from our bank or no more so follow the instruction by sending email to us back with the bank detail details along with your personal details.

Thank you for giving us the opportunity to serve your banking needs. ombes2@gmx.com

Yours sincerely
Board of Directors of Citibank
Sandra Davis
Chief Executive Officer, effective April 16, 2018

Θ = **spam indicator** $= \begin{cases} 1 & \text{IF SPAM} \\ 0 & \text{IF NOT} \end{cases}$

X_1 = **contains** "`million dollars`"

X_2 = **contains** "`Nigerian princess`"

$P(X_1 = 1 \mid \Theta = 1) = 10\%$

$P(X_1 = 1 \mid \Theta = 0) = 3\%$

$P(X_2 = 1 \mid \Theta = 1) = 1\%$

$P(X_2 = 1 \mid \Theta = 0) = 0.01\%$

$P(\Theta = 1) = 20\%$

$P(\Theta = 0) = 80\%$

ASSUME

X_1, X_2 INDEPENDENT GIVEN Θ.

$X_1 = 1$ OBSERVED "MILLION \$"

$X_2 = 0$ DID NOT SEE "NIGERIAN PRINCESS"

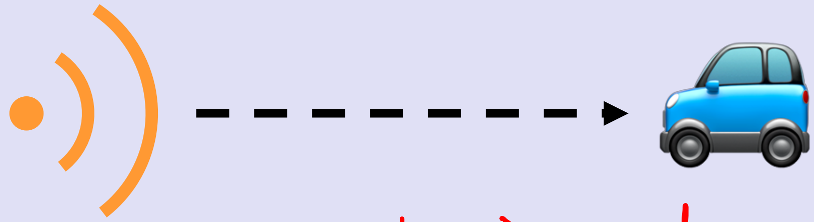
$$P(\theta=1 \mid X_1=1, X_2=0) \propto P(X_1=1, X_2=0 \mid \theta=1)\, P(\theta=1)$$

$$= P(X_1=1 \mid \theta=1)\, P(X_2=0 \mid \theta=1)\, P(\theta=1)$$

$$= 0.1 \cdot 0.99 \cdot 0.2$$

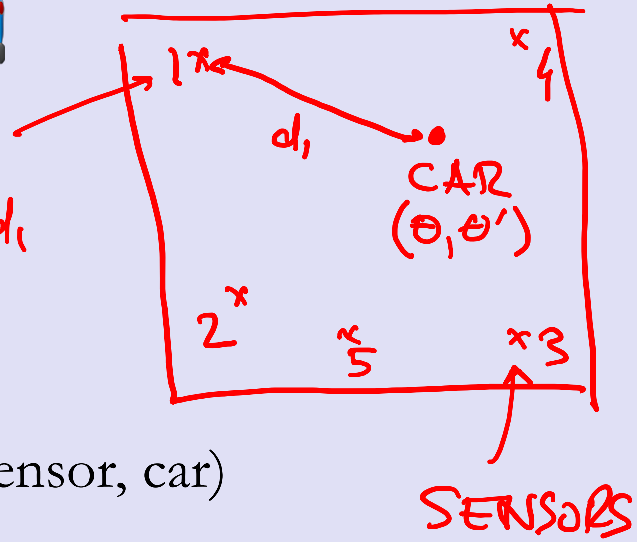
$$P(\theta=0 \mid X_1=1, X_2=0) \propto P(X_1=1 \mid \theta=0)\, P(X_2=0 \mid \theta=0)\, P(\theta=0)$$

$$= 0.03 \cdot 0.9999 \cdot 0.8$$



$P(S_1=1) = e^{-d_1}$

$P(S_1=0) = 1 - e^{-d_1}$



Car position is (θ, θ')

Probability of detection is $e^{-distance(\text{sensor, car})}$

There are 5 **sensors at different positions**

Given that sensors 1, 3, 4 reported detection and 2, 5 didn’t, where is the car?

PROBABILITY MODEL

ASSUME $S_1, \dots, S_5$ ARE INDEPENDENT GIVEN Θ, Θ'

PRIOR: Θ, Θ' INDEPENDENT Normal$(0,1)$

$S_1 = S_3 = S_4 = 1$, $S_2 = S_5 = 0$: WHAT IS Θ, Θ'?

$$f_{\Theta\Theta'}(\theta, \theta' \mid S_1 = S_3 = S_4 = 1, S_2 = S_5 = 0)$$
$$\propto P(S_1 = S_3 = S_4 = 1, S_2 = S_5 = 0 \mid \Theta = \theta, \Theta' = \theta') \cdot f_{\Theta\Theta'}(\theta, \theta')$$
$$= P(S_1 = 1 \mid \Theta = \theta, \Theta' = \theta') P(S_2 = 0 \mid \Theta = \theta, \Theta' = \theta') \dots \cdot f_{\Theta\Theta'}(\theta, \theta')$$
$$= e^{-\sqrt{(x_1-\theta)^2 + (y_1-\theta')^2}} \cdot \left(1 - e^{-\sqrt{(x_2-\theta)^2 + (y_2-\theta')^2}}\right) \cdot \dots \cdot \frac{1}{2\pi} e^{-(\theta^2+\theta'^2)/2}.$$

TO CALCULATE VALUE, DIVIDE BY

$$\iint_{-\infty}^{\infty} e^{-\sqrt{(x_1-\theta)^2 + (y_1-\theta')^2}} \cdot \left(1 - e^{-\sqrt{(x_2-\theta)^2 + (y_2-\theta')^2}}\right) \cdot \dots \cdot \frac{1}{2\pi} e^{-(\theta^2+\theta'^2)/2} \, d\theta \, d\theta'.$$

Point estimation

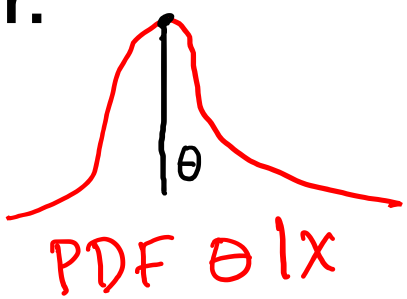
How to turn conditional PDF/PMF $f_{\Theta \mid X}(\theta \mid x)$ estimate into one number?

Conditional expectation (CE) estimator:

$$\mathbf{E}[\theta \mid X = x]$$

Maximum *a posteriori* (MAP) estimator:

$$\operatorname{argmax} f_{\Theta \mid X}(\theta \mid x)$$



Point estimation for normals

$X_i = \text{Normal}(\Theta, 1)$ **independent given** Θ

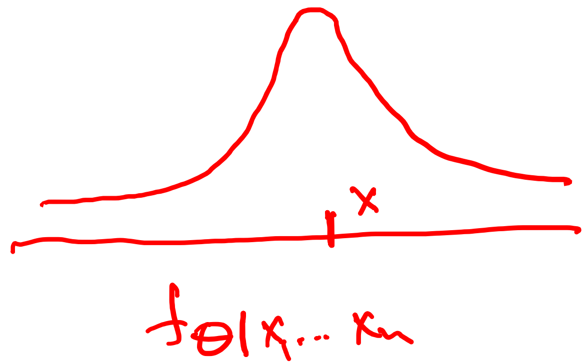
Θ **is** $\text{Normal}(x_0, 1)$

$(\Theta \mid X_1 = x_1, \ldots, X_n = x_n)$ **is** $\text{Normal}(x, 1/\sqrt{n})$

$$x = \frac{x_0 + x_1 + \cdots + x_n}{n+1}$$

CE estimate: $E[\Theta \mid X_1 = x_1 \ldots X_n = x_n] = x$

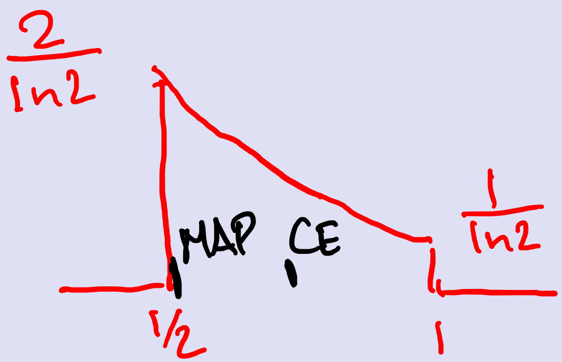
MAP estimate: x



Romeo's model $X = \text{Uniform}(0, \Theta)$

$$\Theta = \text{Uniform}(0, 1)$$

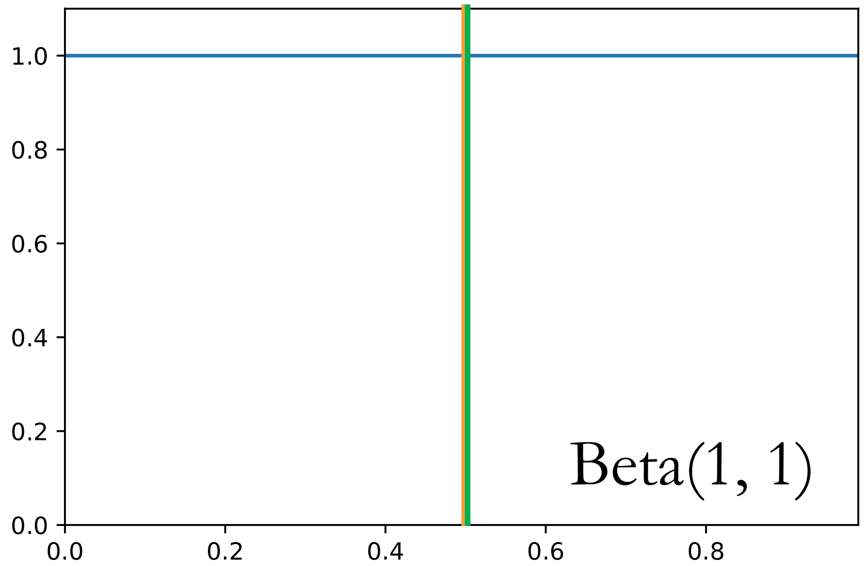
On her first date, Juliet arrives ½ hour late.

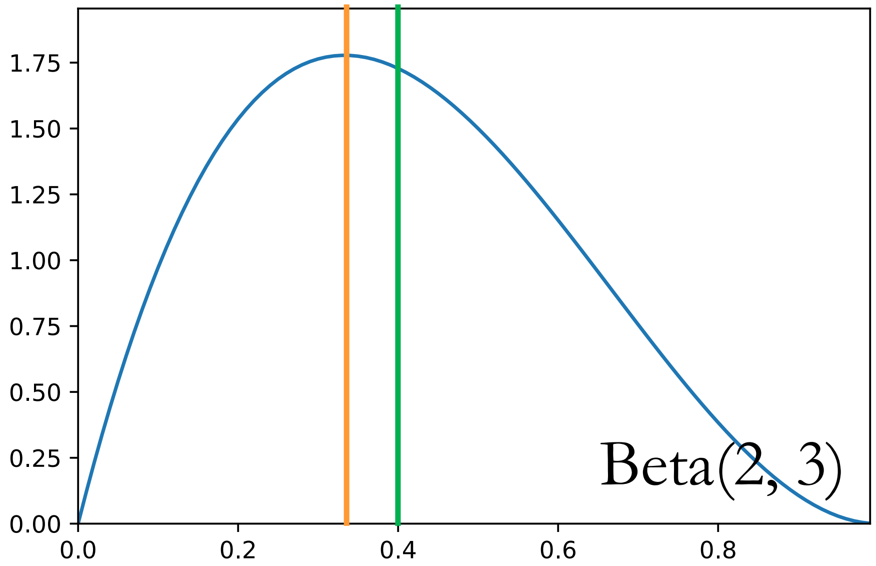


$$f_{\Theta|X}(\theta \mid \tfrac{1}{2}) = \frac{1}{\theta \cdot \ln 2}.$$

CE estimate: $E[\Theta \mid X = \tfrac{1}{2}] = \int_{1/2}^{1} \theta' \frac{1}{\theta' \ln 2} \, d\theta' = \frac{1}{2 \ln 2} \approx 0.72$

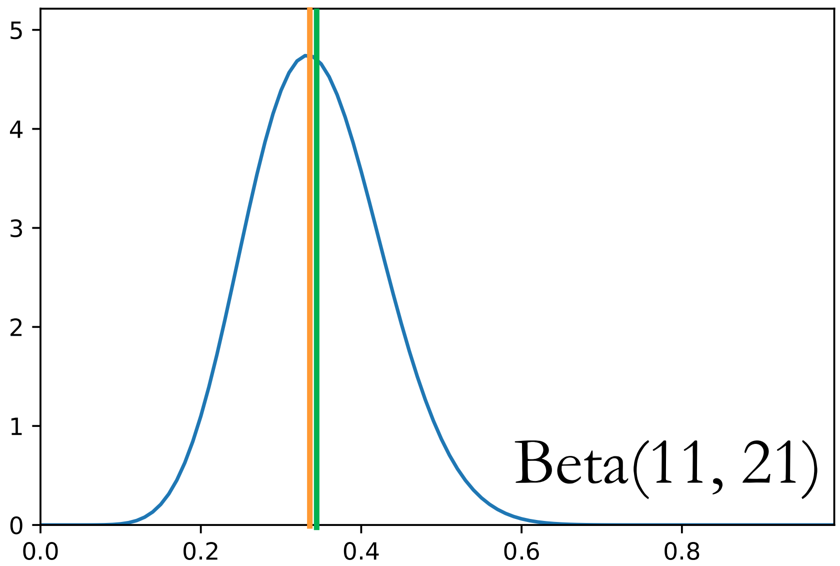
MAP estimate: $\frac{1}{2}$

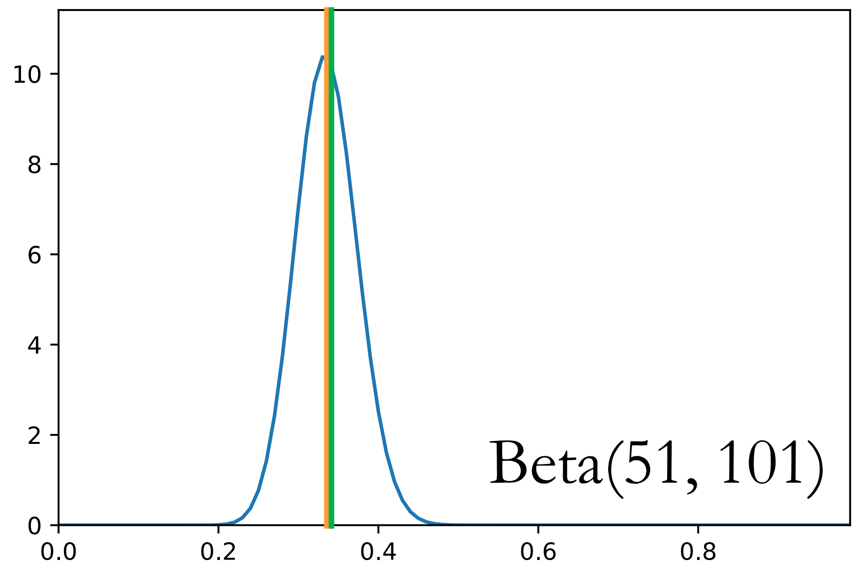




CE $= \alpha/\beta = (k+1)/(n+2)$

MAP $= k/n$





Hypothesis testing

Suppose Θ takes two values (e.g. spam / legit)

$$\textbf{MAP} = \operatorname{argmax} f_{\Theta \mid X}(\theta \mid x)$$

Choose the one for which $f_{\Theta \mid X}(\theta \mid x)$ is larger

$\Theta = 80\%$ **legit,** 20% **spam**

θ	$\mathbf{P}(X_1 \mid \theta)$	$\mathbf{P}(X_2 \mid \theta)$
legit	0.03	0.0001
spam	0.1	0.01

```
The Citibank
concerning wire
transfers of your
fund. Your letter
has been referred to
the (JMCB) Legal
Division for Funds
(US$2.8 Million
Dollars)
```

$P(\text{SPAM} \mid X_1=1, X_2=0) = 0.1 \cdot 0.99 \cdot 0.2 \approx 0.0198$

$P(\text{LEGIT} \mid X_1=1, X_2=0) = 0.03 \cdot 0.9999 \cdot 0.8 \approx 0.0240$

MAP ESTIMATE: LEGIT

Coin A is heads with probability 1/3.

Coin B is tails with probability 1/3.

`HHHT` are 4 flips of a random coin. Which coin was it?

PRIOR: $\Theta = A$ WITH PROB $1/2$
B WITH PROB $1/2$

$$P(A \mid HHHT) \propto P(HHHT \mid A) \cdot P(A)$$
$$= \left(\frac{1}{3}\right)^3 \cdot \frac{2}{3} \cdot \frac{1}{2}$$

$$P(B \mid HHHT) \propto \left(\frac{2}{3}\right)^3 \cdot \frac{1}{3} \cdot \frac{1}{2}$$

$\wedge$ MAP: COIN B.

MAP: MORE HEADS $\rightarrow$ B
MORE TAILS $\rightarrow$ A
2H, 2T $\rightarrow$ A (DOESN'T MATTER)

What is the probability you are wrong, given the outcome is `HHHT`?

W = WRONG DECISION

$$P(W \mid HHHT) = P(\Theta = A \mid HHHT)$$

$$\propto P(HHHT \mid \Theta = A)\, P(\Theta = A)$$

$$= \left(\frac{1}{3}\right)^3 \cdot \frac{2}{3} \cdot \frac{1}{2}$$

$$P(W \mid HHHT) = \frac{\left(\frac{1}{3}\right)^3 \cdot \frac{2}{3} \cdot \frac{1}{2}}{\left(\frac{1}{3}\right)^3 \cdot \frac{2}{3} \cdot \frac{1}{2} + \left(\frac{2}{3}\right)^3 \cdot \frac{1}{3} \cdot \frac{1}{2}} = \frac{1}{1+4} = 20\%$$

What is the probability you are wrong on average?

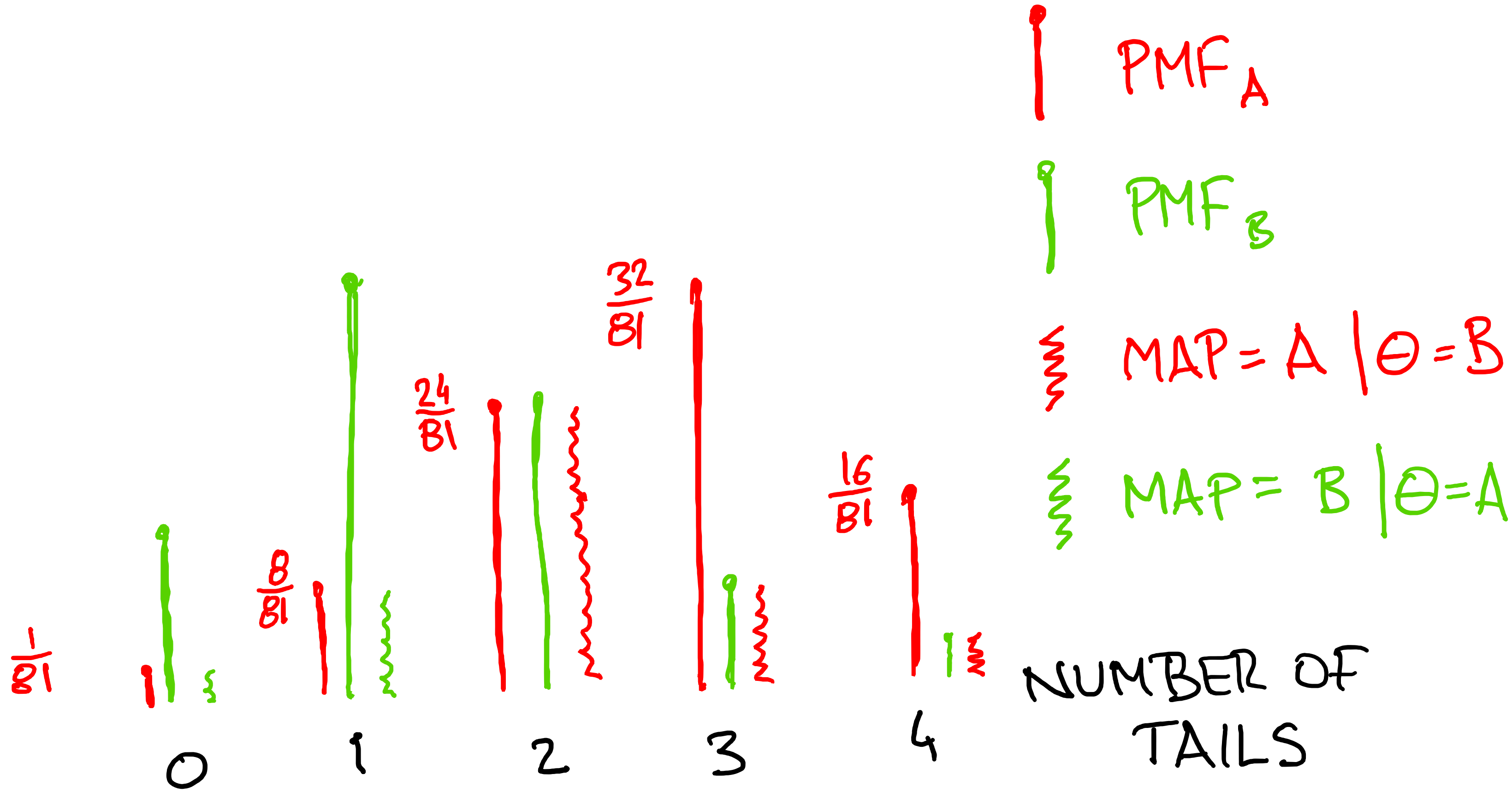
$$P(W) = P(MAP = A, \Theta = B) + P(MAP = B, \Theta = A)$$

$$= \frac{1}{2} \cdot P(MAP = A \mid \Theta = B) + \frac{1}{2} P(MAP = B \mid \Theta = A)$$

$$= \frac{1}{2}\left(\left(\frac{1}{3}\right)^4 + 4 \cdot \frac{2}{3} \cdot \left(\frac{1}{3}\right)^3 + 6 \cdot \left(\frac{2}{3}\right)^2 \left(\frac{1}{3}\right)^2\right) + \frac{1}{2}\left(\left(\frac{1}{3}\right)^4 + 4 \cdot \left(\frac{1}{3}\right)^3 \cdot \frac{2}{3}\right)$$

$$= \frac{21}{81} \approx 26\%$$

Hypothesis testing error



An car-jack detector outputs $\mathrm{Normal}(0, \frac{1}{2})$ if there is no intruder and $\mathrm{Normal}(1, \frac{1}{2})$ if there is. When should alarm activate?