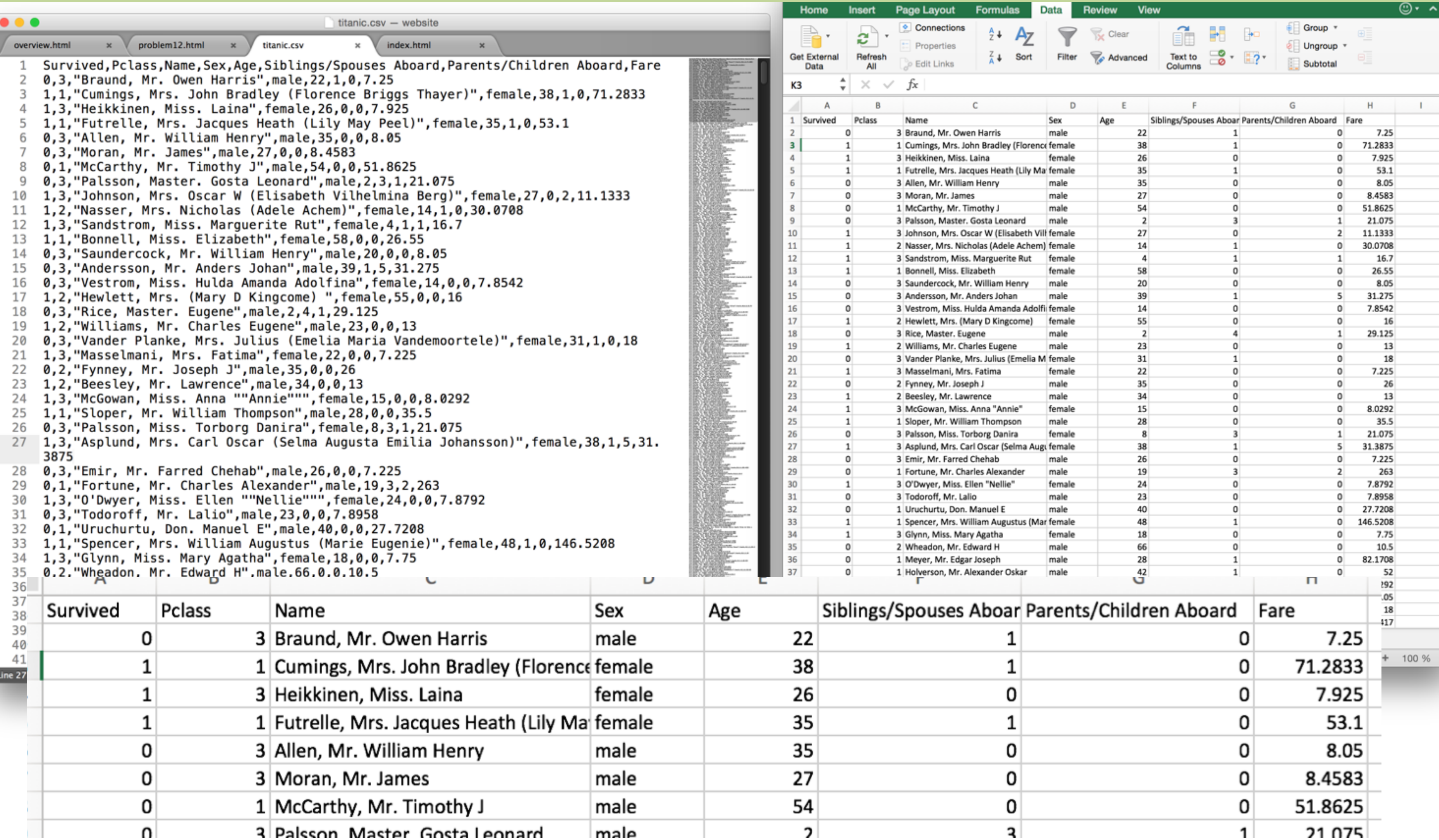
Properties of Joint Distributions II

Chris Piech
CS109, Stanford University

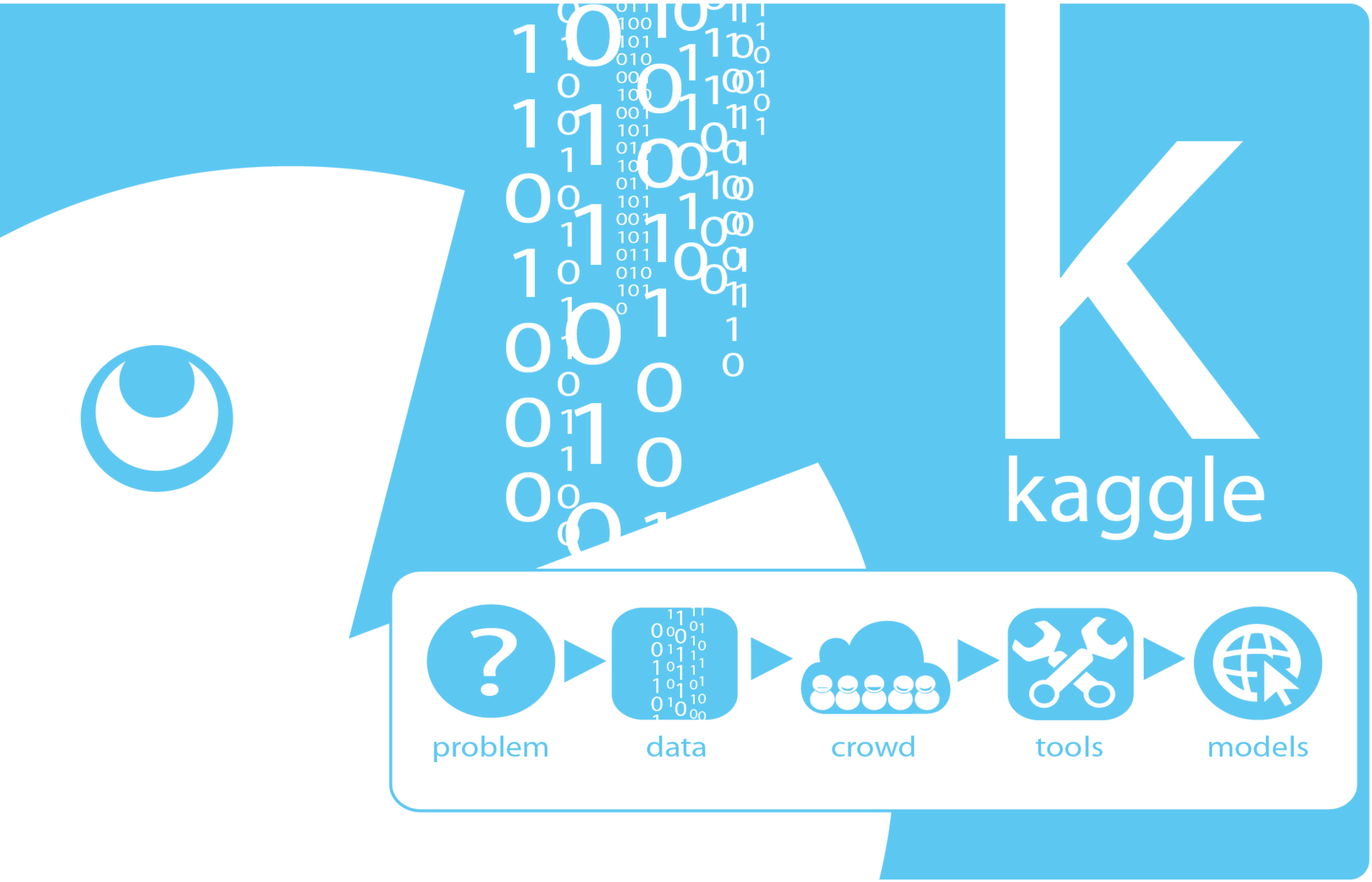

TITANIC

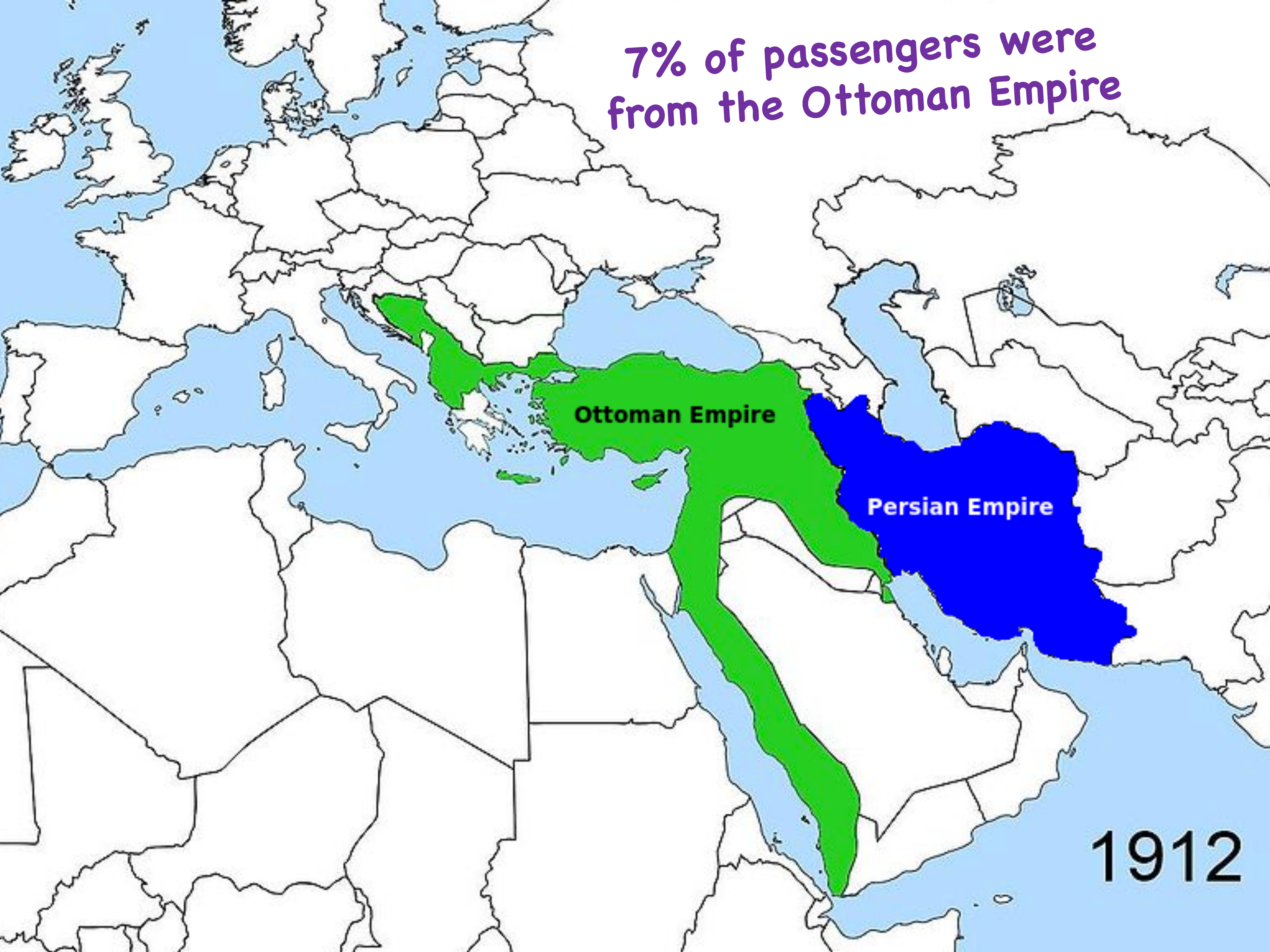
Titanic Probability




```
Survived,Pclass,Name,Sex,Age,Siblings/Spouses Aboard,Parents/Children Aboard,Fare
0,3,"Braund, Mr. Owen Harris",male,22,1,0,7.25
1,1,"Cumings, Mrs. John Bradley (Florence Briggs Thayer)",female,38,1,0,71.2833
1,3,"Heikkinen, Miss. Laina",female,26,0,0,7.925
1,1,"Futrelle, Mrs. Jacques Heath (Lily May Peel)",female,35,1,0,53.1
0,3,"Allen, Mr. William Henry",male,35,0,0,8.05
0,3,"Moran, Mr. James",male,27,0,0,8.4583
0,1,"McCarthy, Mr. Timothy J",male,54,0,0,51.8625
0,3,"Palsson, Master. Gosta Leonard",male,2,3,1,21.075
1,3,"Johnson, Mrs. Oscar W (Elisabeth Vilhelmina Berg)",female,27,0,2,11.1333
1,2,"Nasser, Mrs. Nicholas (Adele Achem)",female,14,1,0,30.0708
1,3,"Sandstrom, Miss. Marguerite Rut",female,4,1,1,16.7
1,1,"Bonnell, Miss. Elizabeth",female,58,0,0,26.55
0,3,"Saundercock, Mr. William Henry",male,20,0,0,8.05
0,3,"Andersson, Mr. Anders Johan",male,39,1,5,31.275
0,3,"Vestrom, Miss. Hulda Amanda Adolfina",female,14,0,0,7.8542
1,2,"Hewlett, Mrs. (Mary D Kingcome) ",female,55,0,0,16
0,3,"Rice, Master. Eugene",male,2,4,1,29.125
1,2,"Williams, Mr. Charles Eugene",male,23,0,0,13
0,3,"Vander Planke, Mrs. Julius (Emelia Maria Vandemoortele)",female,31,1,0,18
1,3,"Masselmani, Mrs. Fatima",female,22,0,0,7.225
0,2,"Fynney, Mr. Joseph J",male,35,0,0,26
1,2,"Beesley, Mr. Lawrence",male,34,0,0,13
1,3,"McGowan, Miss. Anna ""Annie""",female,15,0,0,8.0292
1,1,"Sloper, Mr. William Thompson",male,28,0,0,35.5
0,3,"Palsson, Miss. Torborg Danira",female,8,3,1,21.075
1,3,"Asplund, Mrs. Carl Oscar (Selma Augusta Emilia Johansson)",female,38,1,5,31.
3875
0,3,"Emir, Mr. Farred Chehab",male,26,0,0,7.225
0,1,"Fortune, Mr. Charles Alexander",male,19,3,2,263
1,3,"O'Dwyer, Miss. Ellen ""Nellie""",female,24,0,0,7.8792
0,3,"Todoroff, Mr. Lalio",male,23,0,0,7.8958
0,1,"Uruchurtu, Don. Manuel E",male,40,0,0,27.7208
1,1,"Spencer, Mrs. William Augustus (Marie Eugenie)",female,48,1,0,146.5208
1,3,"Glynn, Miss. Mary Agatha",female,18,0,0,7.75
0,2,"Wheadon, Mr. Edward H",male,66,0,0,10.5
```

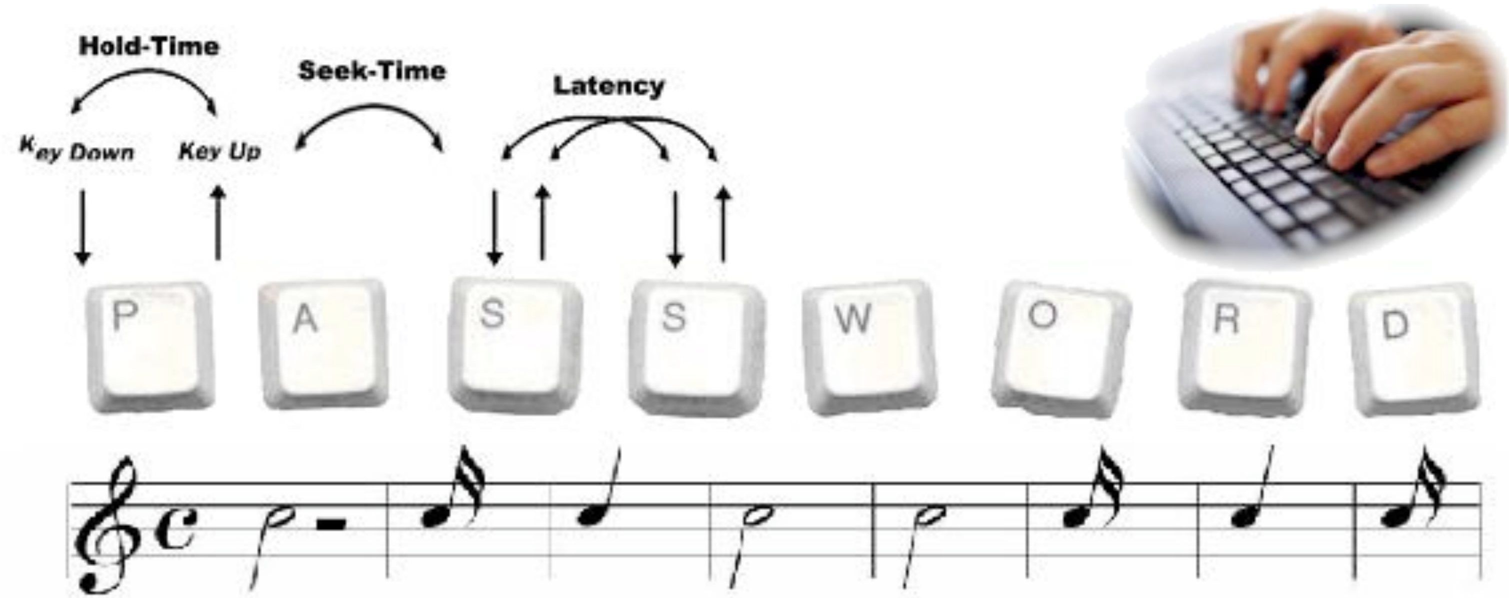
Survived	Pclass	Name	Sex	Age	Siblings/Spouses Aboar	Parents/Children Aboard	Fare
0	3	Braund, Mr. Owen Harris	male	22	1	0	7.25
1	1	Cumings, Mrs. John Bradley (Florence	female	38	1	0	71.2833
1	3	Heikkinen, Miss. Laina	female	26	0	0	7.925
1	1	Futrelle, Mrs. Jacques Heath (Lily Ma	female	35	1	0	53.1
0	3	Allen, Mr. William Henry	male	35	0	0	8.05
0	3	Moran, Mr. James	male	27	0	0	8.4583
0	1	McCarthy, Mr. Timothy J	male	54	0	0	51.8625
0	3	Palsson, Master. Gosta Leonard	male	2	3	1	21.075


kaggle
problem
data
crowd
tools
models


7% of passengers were
from the Ottoman Empire
Ottoman Empire
Persian Empire
1912

Europe
Kiev
Paris
France
Vienna
Budapest
Jassy
Christian vassal states
Venice
Belgrade
Bucharest
Caffa
Marseille
Italy
Varna
Black Sea
Caspian Sea
Barcelona
Rome
Prizren
Sofia
Constantinople
Baku
Salonica
Naples
Spain
Sicily
Athens
Tunis
Alexandretta
Algiers
Algeria
Rhodes
Mediterranean Sea
Damascus
Baghdad
Tripoli
Alexandria
Jerusalem
Cairo
Lybia
Egypt
750 miles
1000 km
Arabia
Africa
Medina
Mecca
Red Sea
Kassala
Zeila
THE OTTOMAN EMPIRE IN 1683
Ottoman Beylik, 1300
Acquisitions, 1300 - 1359
Acquisitions, 1359 - 1451
Acquisitions, 1451 - 1481 (Mehmed II)
Acquisitions, 1512 - 1520 (Selim I)
Acquisitions, 1520 - 1566 (Suleiman the Magnificent)
Acquisitions, 1566 - 1683

Biometric Keystrokes

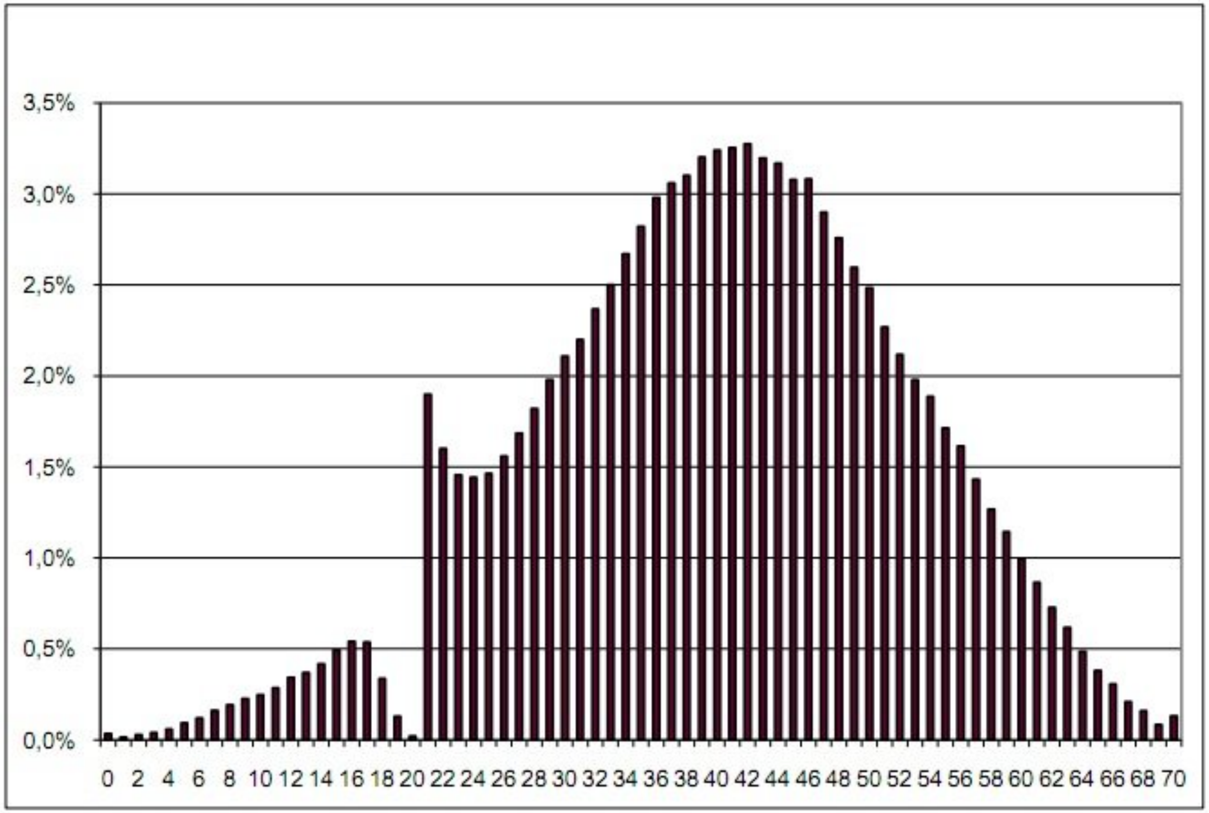


Altruism?

Scores for a standardized test that students in Poland are required to pass before moving on in school

See if you can guess the minimum score to pass the test.

2.1. Poziom podstawowy



Wykres 1. Rozkład wyników na poziomie podstawowym

http://freakonomics.com/2011/07/07/another-case-of-teacher-cheating-or-is-it-just-altruism/comment-page-2/

Boolean Operation on Variable = Event

Recall: any boolean question about a random variable makes for an event. For example:

$$P(X \leq 5)$$

$$P(Y = 6)$$

$$P(5 \leq Z \leq 10)$$

Joint Random Variables

- [x] Use a joint table, density function or CDF to solve probability question
- [x] Use and find **independence** of random variables
- [x] Think about **conditional** probabilities with joint variables (which might be continuous)
- [ ] What happens when you **add** random variables?

Independence and Random Variables

Independent Continuous Variables

- Two continuous random variables X and Y are called **independent** if:

  P(X ≤ *a*, Y ≤ *b*) = P(X ≤ *a*) P(Y ≤ *b*) for any *a*, *b*

- Equivalently:

$$F_{X,Y}(a,b) = F_X(a)F_Y(b) \quad \text{for all } a,b$$

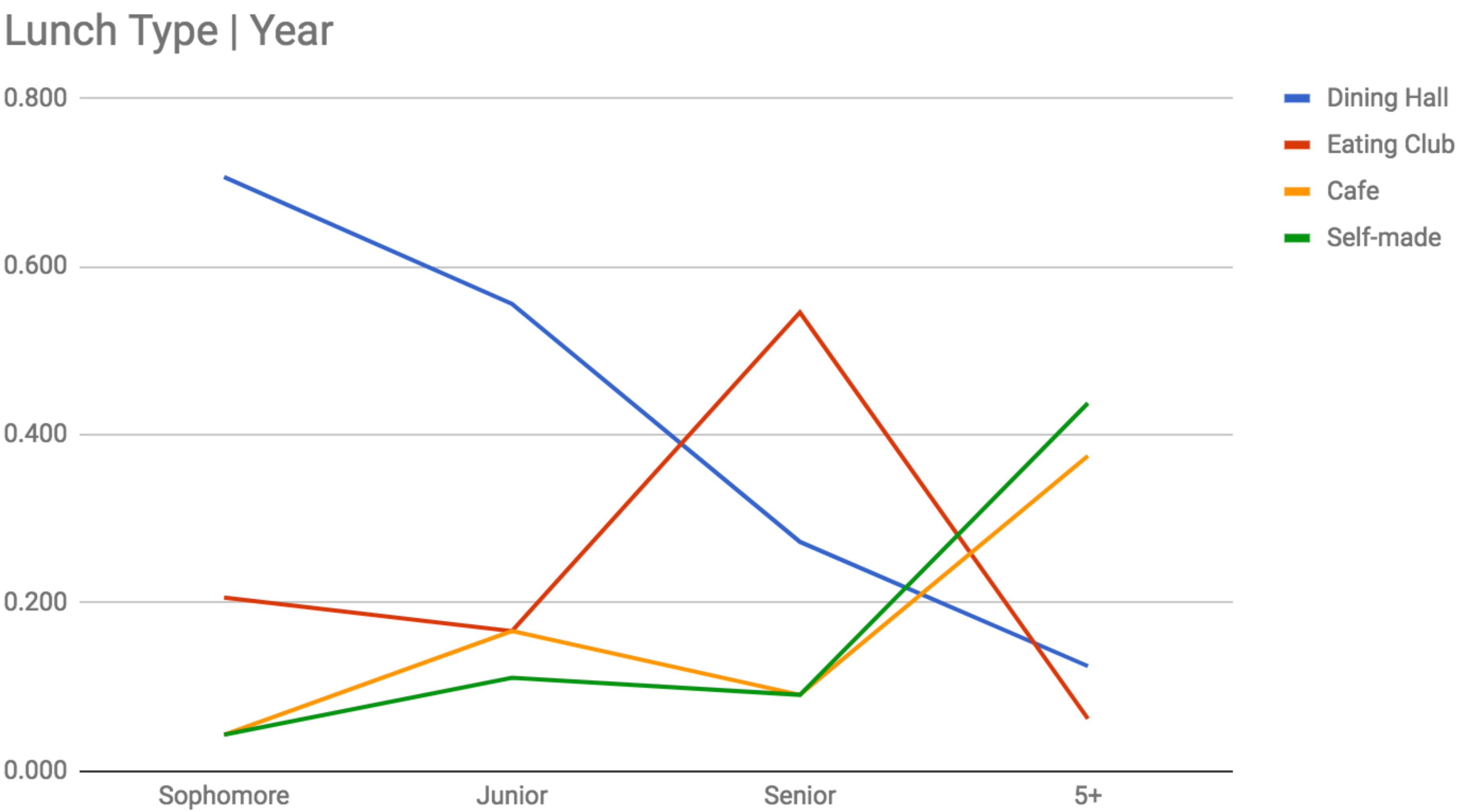
$$f_{X,Y}(a,b) = f_X(a)f_Y(b) \quad \text{for all } a,b$$

- More generally, joint density factors separately:

$$f_{X,Y}(x,y) = h(x)g(y) \quad \text{where } -\infty < x, y < \infty$$

Conditionals with multiple variables

Lunch | Year



Lunch Type | Year
0.800
0.600
0.400
0.200
0.000
Sophomore
Junior
Senior
5+
Dining Hall
Eating Club
Cafe
Self-made

Continuous Conditional Distributions

Let X and Y be continuous random variables

$$P(X = x | Y = y) = \frac{P(X = x, Y = y)}{P(Y = y)}$$

$$f_{X|Y}(x|y) \cdot \epsilon_x = \frac{f_{X,Y}(x, y) \cdot \epsilon_x \cdot \epsilon_y}{f_Y(y) \cdot \epsilon_y}$$

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

Bayes with a mix

Let X be a continuous random variable

Let N be a discrete random variable

$$P(X = x | N = n) = \frac{P(N = n | X = x) P(X = x)}{P(N = n)}$$

$$P_{X|N}(x|n) = \frac{P_{N|X}(n|x) P_X(x)}{P_N(n)}$$

$$f_{X|N}(x|n) \cdot \epsilon_x = \frac{P_{N|X}(n|x) f_X(x) \cdot \epsilon_x}{P_N(n)}$$

$$\boxed{f_{X|N}(x|n) = \frac{P_{N|X}(n|x) f_X(x)}{P_N(n)}}$$

All the Bayes Belong to Us

M,N are discrete. X, Y are continuous

OG Bayes

$$p_{M|N}(m|n) = \frac{P_{N|M}(n|m)p_M(m)}{p_N(n)}$$

Mix Bayes #1

$$f_{X|N}(x|n) = \frac{P_{N|X}(n|x)f_X(x)}{P_N(n)}$$

Mix Bayes #2

$$p_{N|X}(n|x) = \frac{f_{X|N}(x|n)p_N(n)}{f_X(x)}$$

All Continuous

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

Let's Do an Example

- X and Y are continuous RVs with PDF:

$$f(x,y)=\begin{cases}\frac{12}{5}x(2-x-y) & \text{where } 0<x,y<1\\ 0 & \text{otherwise}\end{cases}$$

  - Compute conditional density: $f_{X|Y}(x\mid y)$

$$f_{X|Y}(x\mid y)=\frac{f_{X,Y}(x,y)}{f_Y(y)}=\frac{f_{X,Y}(x,y)}{\int_0^1 f_{X,Y}(x,y)\,dx}$$

$$=\frac{\frac{12}{5}x(2-x-y)}{\int_0^1 \frac{12}{5}x(2-x-y)\,dx}=\frac{x(2-x-y)}{\int_0^1 x(2-x-y)\,dx}=\frac{x(2-x-y)}{\left[x^2-\frac{x^3}{3}-\frac{x^2y}{2}\right]_0^1}$$

$$=\frac{x(2-x-y)}{\frac{2}{3}-\frac{y}{2}}=\frac{6x(2-x-y)}{4-3y}$$

Warmup: Bayes Revisited

Likelihood of evidence

Posterior belief

Prior belief

$$P(B|E) = \frac{P(E|B) \quad P(B)}{P(E)}$$

Normalization constant

Warmup: Bivariate Normal

- X, Y follow a symmetric bivariate normal distribution if they have joint PDF:

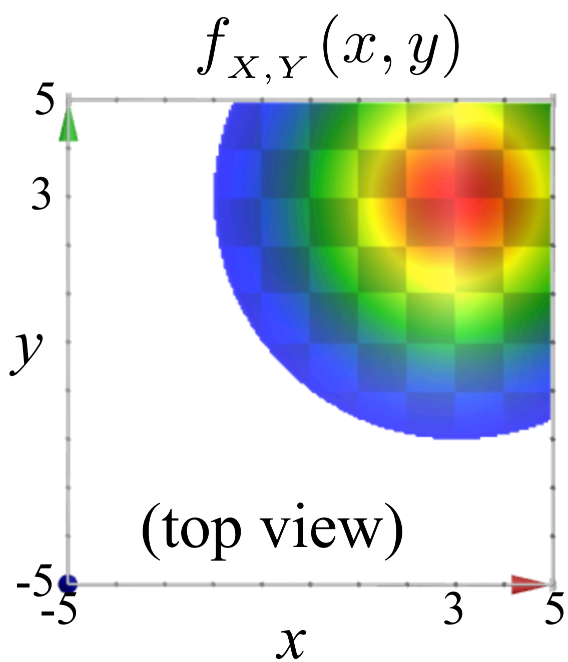
$$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{[(x-\mu_x)^2+(y-\mu_y)^2]}{2\cdot\sigma^2}}$$

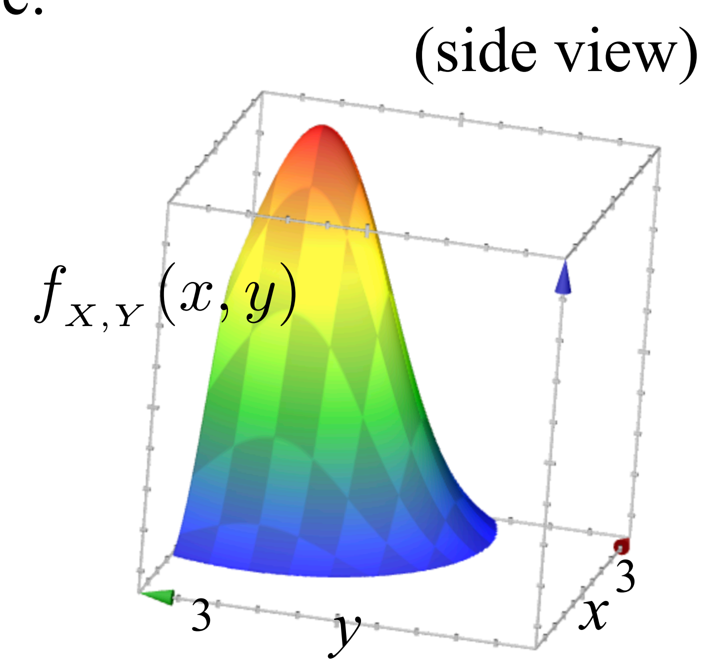
Here is an example where:

$$\mu_x = 3$$

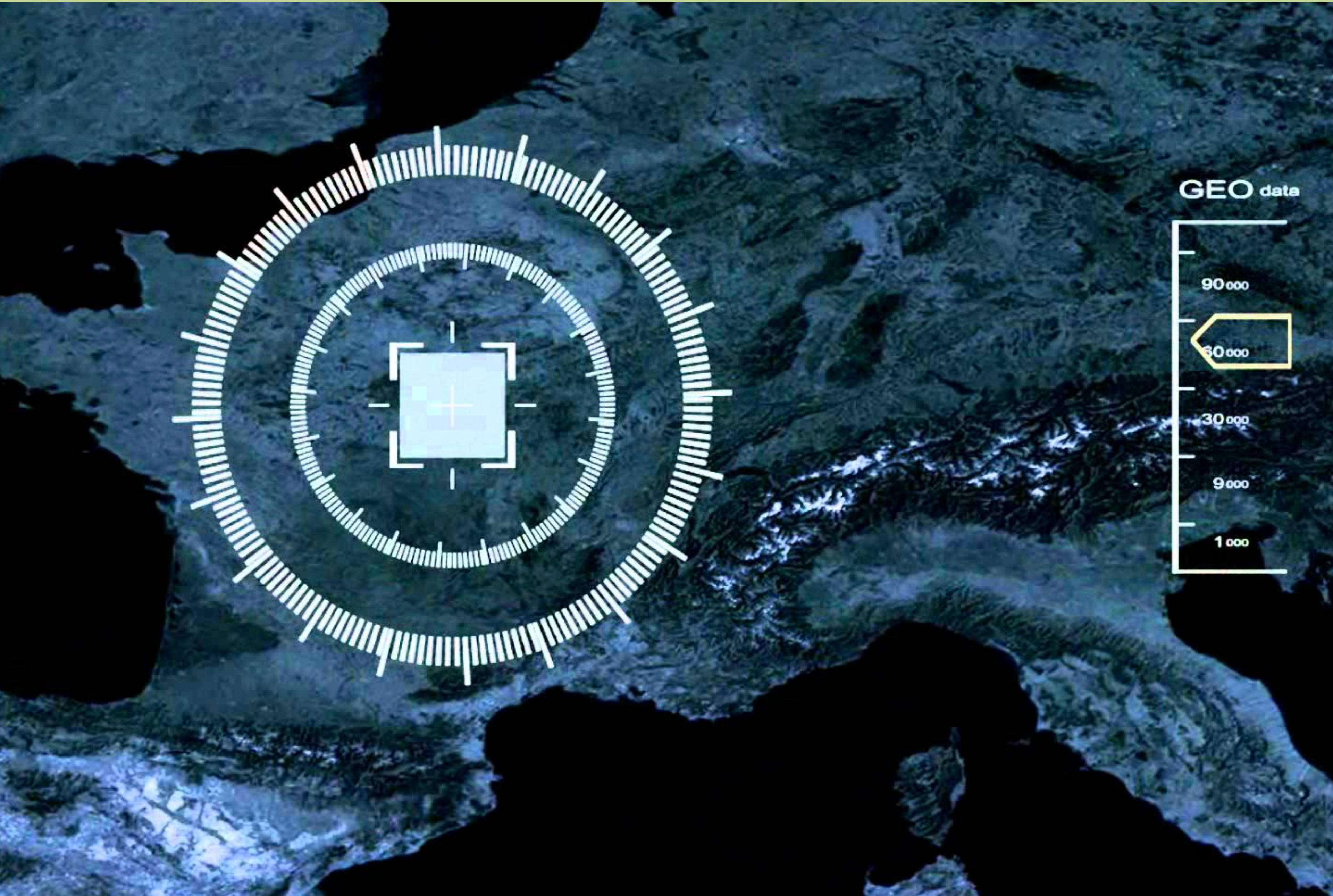
$$\mu_y = 3$$

$$\sigma = 2$$



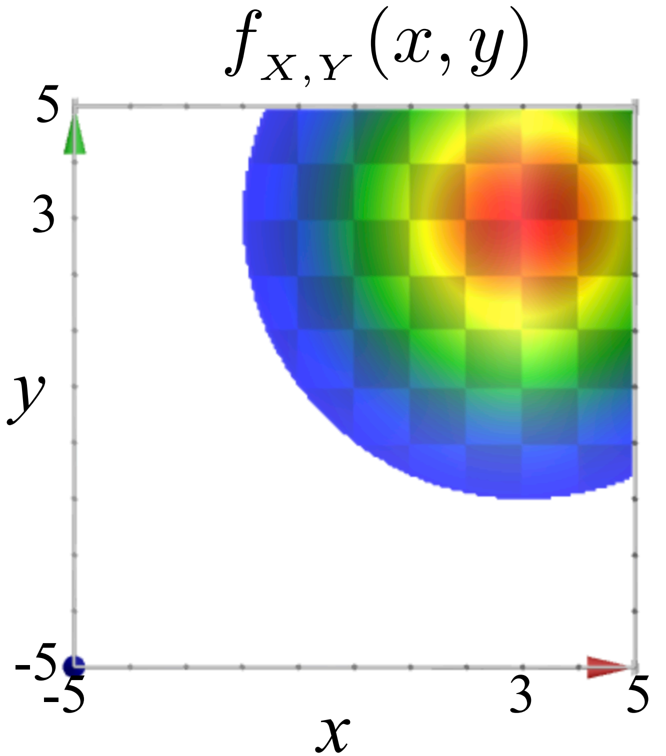


Tracking in 2D Space?



Tracking in 2D Space: Prior

Prior belief: $$f_{X,Y}(x,y) = \frac{1}{2\pi\sigma^2} \cdot e^{-\frac{[(x-\mu_x)^2+(y-\mu_y)^2]}{2\cdot\sigma^2}}$$

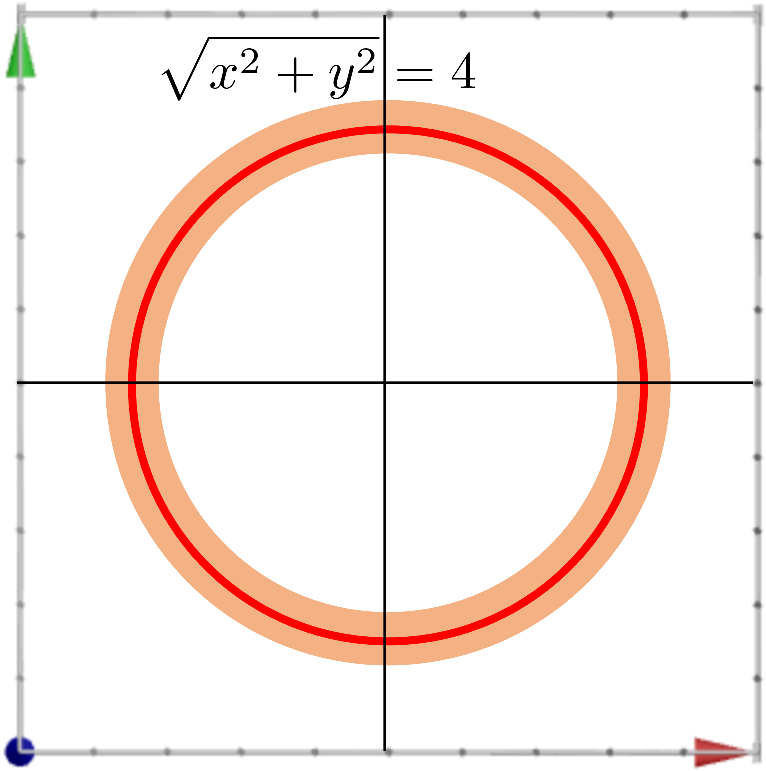


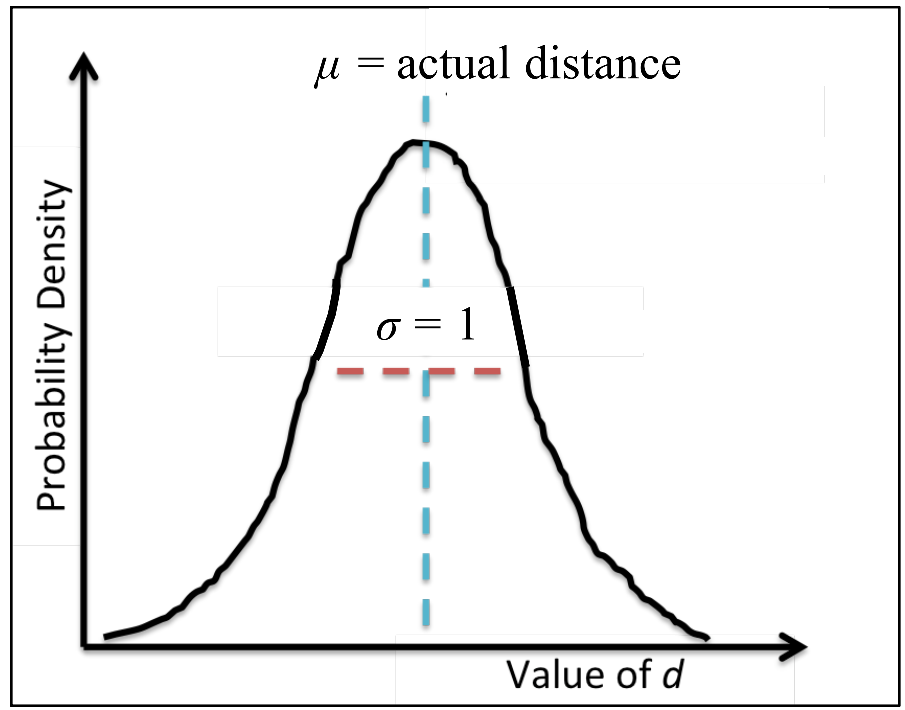
$$\mu_x = 3$$
$$\mu_y = 3$$
$$\sigma = 2$$

Prior belief with K: $$f_{X,Y}(x,y) = K \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}$$

Tracking in 2D Space: Observation!

Observe a ping of the object that is distance $D = 4$ away!



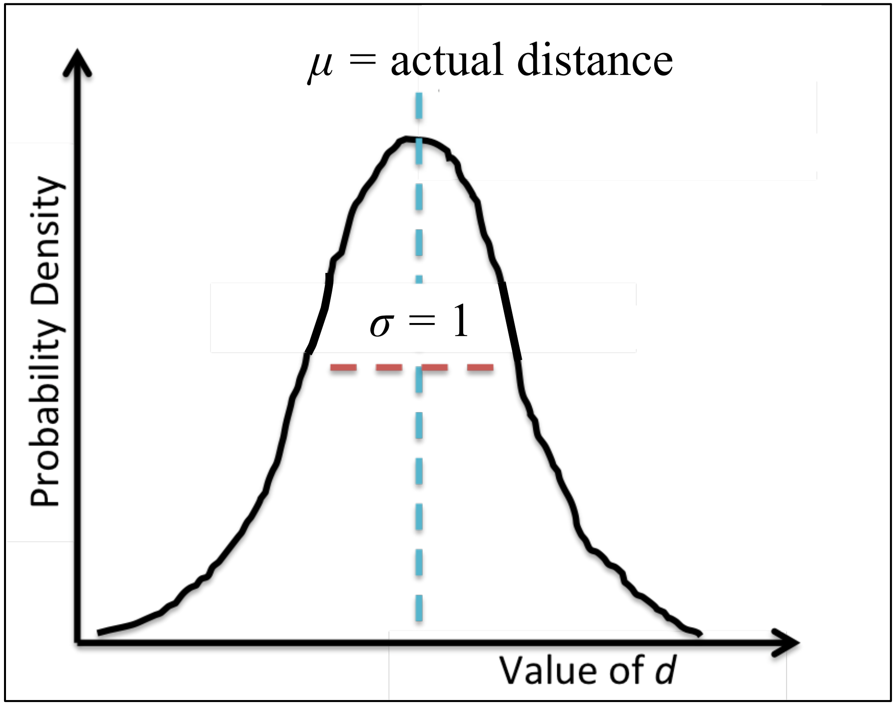


Know that the distance of a ping is normal with respect to the true distance

Tracking in 2D Space: Observation!

Observe a ping of the object that is distance $D = 4$ away!

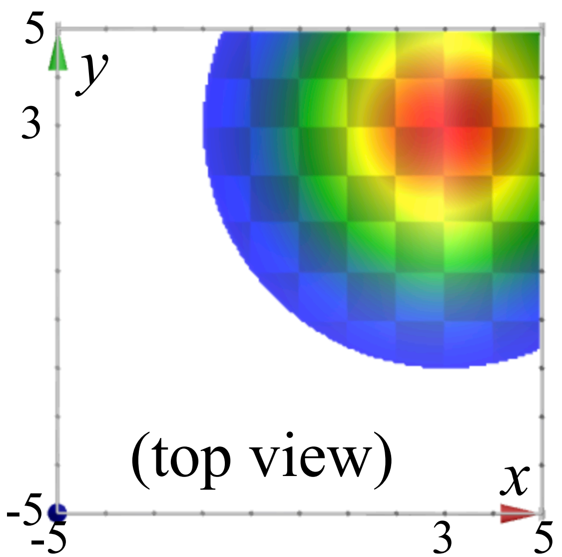
$$D|X,Y \sim N(\mu = \sqrt{x^2 + y^2}, \sigma^2 = 1)$$



Know that the distance of a ping is normal with respect to the true distance. What is the PDF of D?

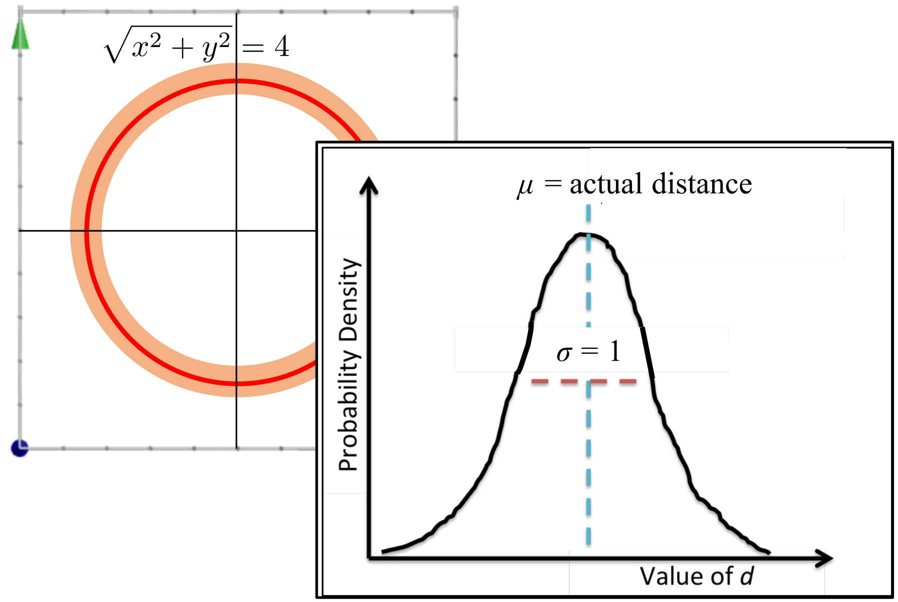
Tracking in 2D Space: New Belief

$$f(X=x, Y=y) = K_1 \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}$$



Prior

Observation



$$f(D=d|X=x, Y=y) = K_2 \cdot e^{-\frac{[d-\sqrt{x^2+y^2}]^2}{2}}$$

What is your ***new*** belief for the location of the object being tracked? Your joint probability density function can be expressed with a constant

Tracking in 2D Space: New Belief

$$f(X=x,Y=y|D=4) = \frac{f(D=4|X=x,Y=y)\cdot f(X=x,Y=y)}{f(D=4)}$$

$$= \frac{K_1 \cdot e^{-\frac{[4-\sqrt{x^2+y^2})^2]}{2}} \cdot K_2 \cdot e^{-\frac{[(x-3)^2+(y-3)^2]}{8}}}{f(D=4)}$$

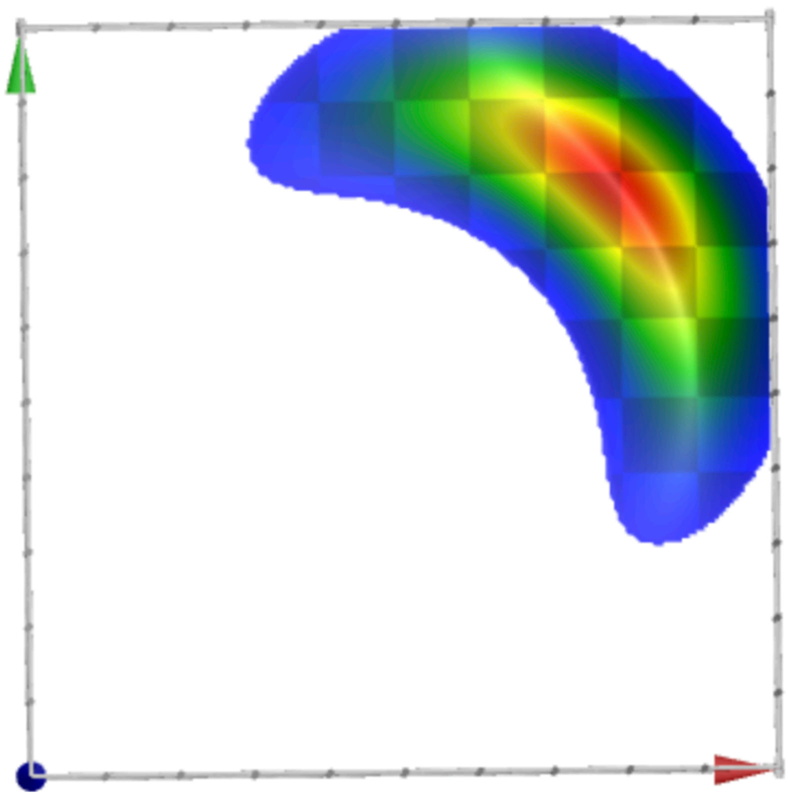
$$= \frac{K_3 \cdot e^{-\left[\frac{[4-\sqrt{x^2+y^2})^2]}{2}+\frac{[(x-3)^2+(y-3)^2]}{8}\right]}}{f(D=4)}$$

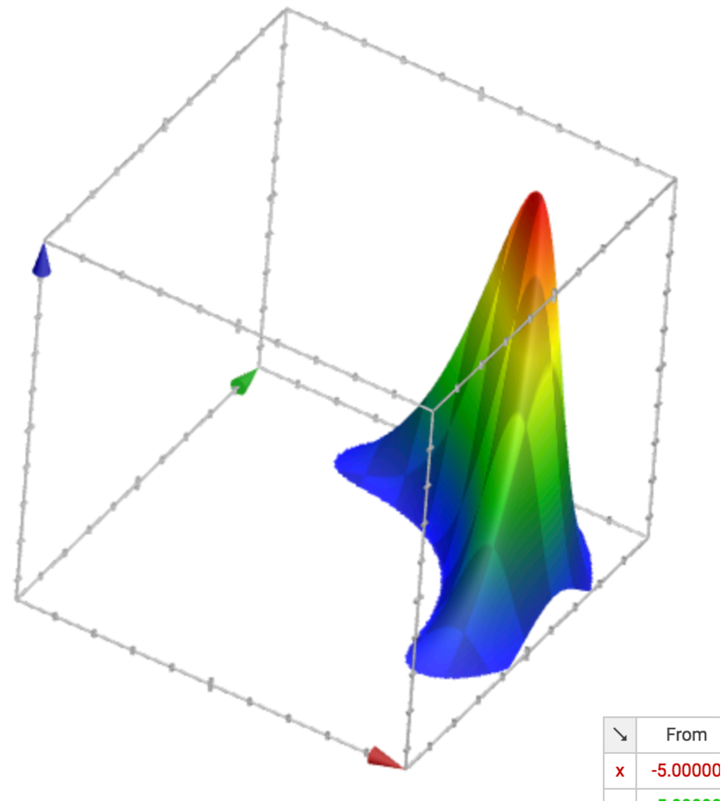
$$= K_4 \cdot e^{-\left[\frac{(4-\sqrt{x^2+y^2})^2}{2}+\frac{[(x-3)^2+(y-3)^2]}{8}\right]}$$

For your notes…

Tracking in 2D Space: Posterior

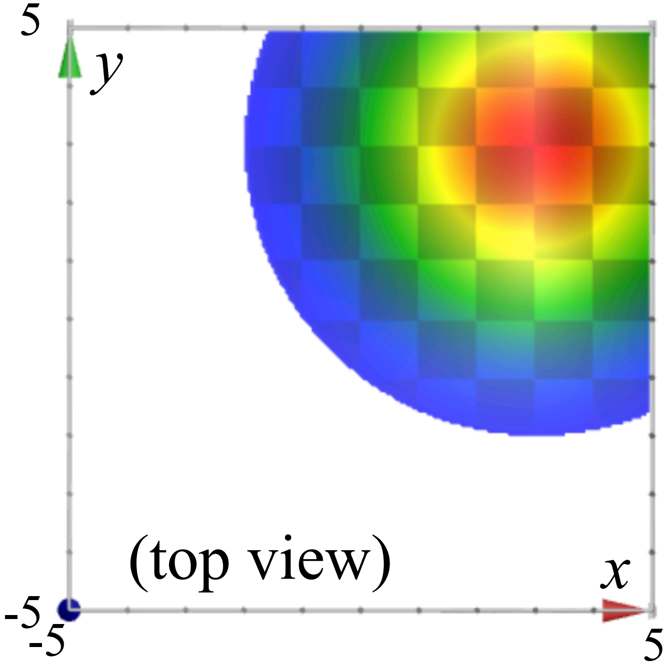
$$f_{X,Y|D}(x,y|4) = K_4 \cdot e^{-\left[\frac{(4-\sqrt{x^2+y^2})^2}{2} + \frac{[(x-3)^2+(y-3)^2]}{8}\right]}$$





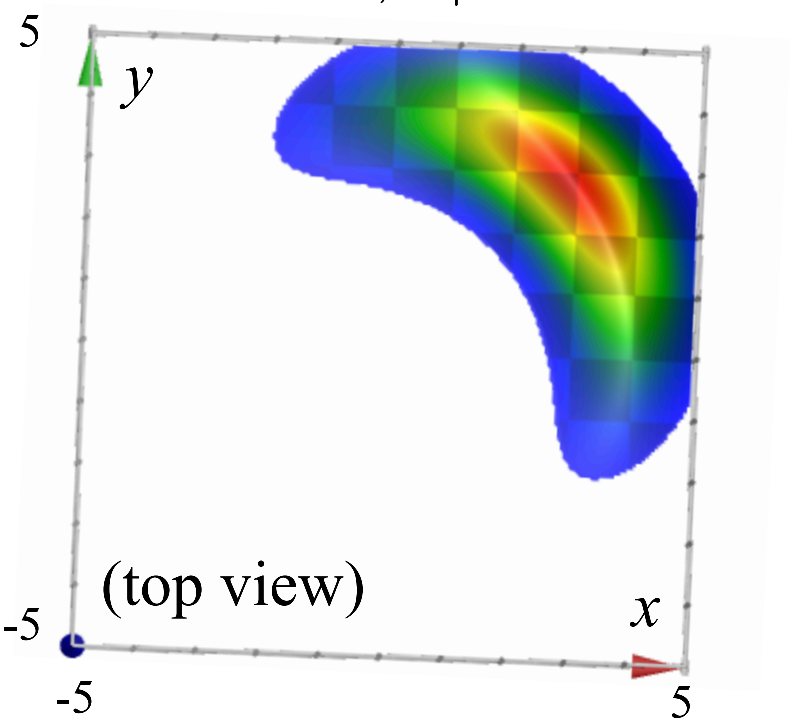
Tracking in 2D Space: Posterior

$f_{X,Y}$



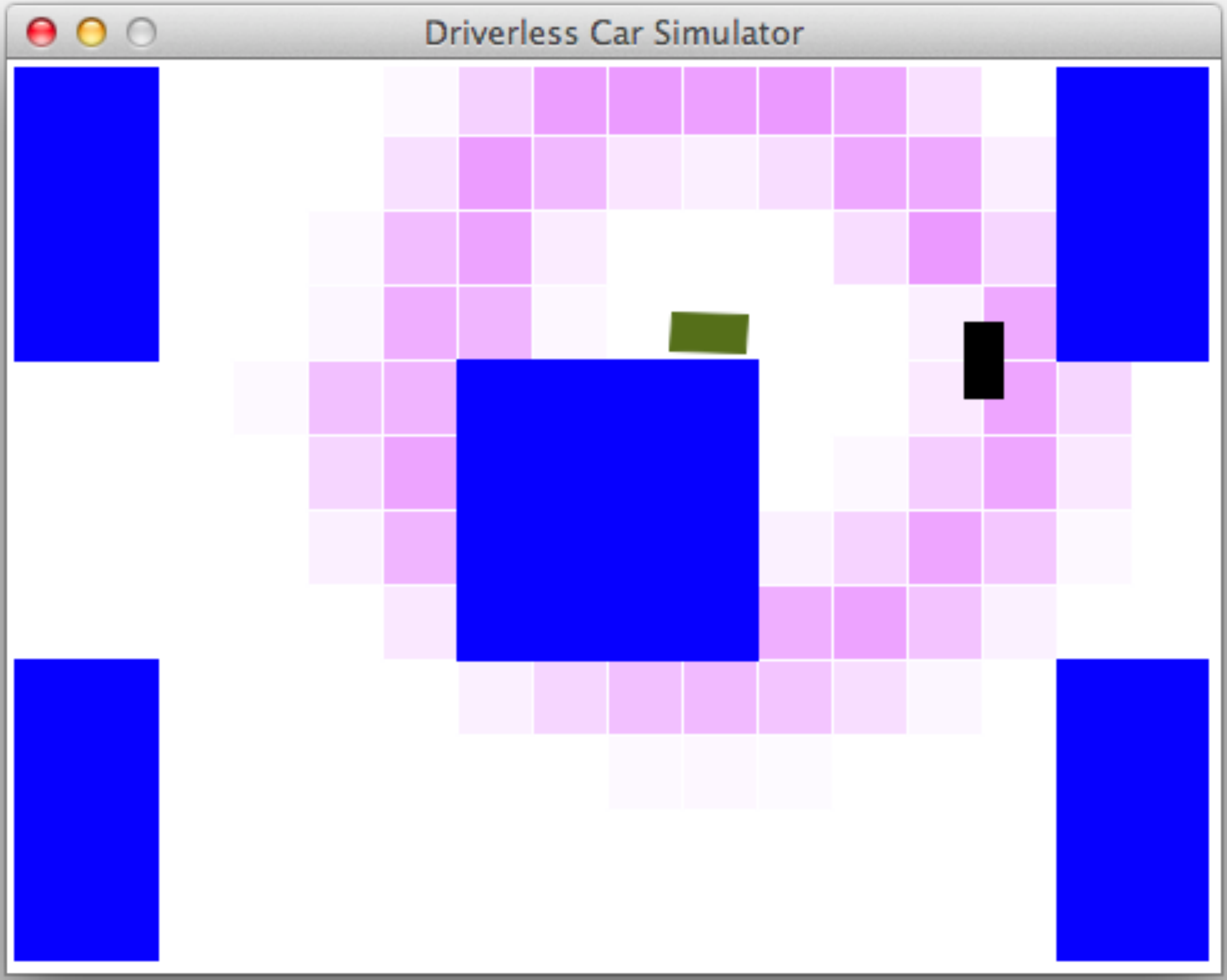
Prior

$f_{X,Y|D}$



Posterior

Tracking in 2D Space: CS221



What happens when you add random variables?

Sum of Independent Binomials

- Let X and Y be independent random variables
  - $X \sim Bin(n_1, p)$ and $Y \sim Bin(n_2, p)$
  - $X + Y \sim Bin(n_1 + n_2, p)$
- Intuition:
  - X has n_1 trials and Y has n_2 trials
    - Each trial has same "success" probability p
  - Define Z to be $n_1 + n_2$ trials, each with success prob. p
  - $Z \sim Bin(n_1 + n_2, p)$, and also $Z = X + Y$

Sum of Independent Poissons

First! Recall the Binomial Theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$$

Sum of Independent Poissons

- Let X and Y be independent random variables
  - X ~ Poi(λ_1) and Y ~ Poi(λ_2)
  - X + Y ~ Poi(λ_1 + λ_2)
- Proof: (just for reference)
  - Rewrite (X + Y = *n*) as (X = *k*, Y = *n* – *k*) where 0 ≤ *k* ≤ *n*

$$P(X+Y=n)=\sum_{k=0}^{n}P(X=k,Y=n-k)=\sum_{k=0}^{n}P(X=k)P(Y=n-k)$$

$$=\sum_{k=0}^{n}e^{-\lambda_1}\frac{\lambda_1^k}{k!}e^{-\lambda_2}\frac{\lambda_2^{n-k}}{(n-k)!}=e^{-(\lambda_1+\lambda_2)}\sum_{k=0}^{n}\frac{\lambda_1^k\lambda_2^{n-k}}{k!(n-k)!}=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k}$$

  - Noting Binomial theorem: $(\lambda_1+\lambda_2)^n=\sum_{k=0}^{n}\frac{n!}{k!(n-k)!}\lambda_1^k\lambda_2^{n-k}$
  - $P(X+Y=n)=\frac{e^{-(\lambda_1+\lambda_2)}}{n!}\left(\lambda_1+\lambda_2\right)^n$ so, X + Y = n ~ Poi(λ_1 + λ_2)

Reference: Sum of Independent RVs

- Let X and Y be independent Binomial RVs
  - $X \sim \text{Bin}(n_1, p)$ and $Y \sim \text{Bin}(n_2, p)$
  - $X + Y \sim \text{Bin}(n_1 + n_2, p)$
  - More generally, let $X_i \sim \text{Bin}(n_i, p)$ for $1 \leq i \leq N$, then

$$\left(\sum_{i=1}^{N} X_i\right) \sim \text{Bin}\left(\sum_{i=1}^{N} n_i,\ p\right)$$

- Let X and Y be independent Poisson RVs
  - $X \sim \text{Poi}(\lambda_1)$ and $Y \sim \text{Poi}(\lambda_2)$
  - $X + Y \sim \text{Poi}(\lambda_1 + \lambda_2)$
  - More generally, let $X_i \sim \text{Poi}(\lambda_i)$ for $1 \leq i \leq N$, then

$$\left(\sum_{i=1}^{N} X_i\right) \sim \text{Poi}\left(\sum_{i=1}^{N} \lambda_i\right)$$

If only it were always that simple

Convolution of Probability Distributions

We talked about sum of Binomial and Poisson…who's missing from this party?

Uniform.

Summation: not just for the 1%

Dance, Dance Convolution

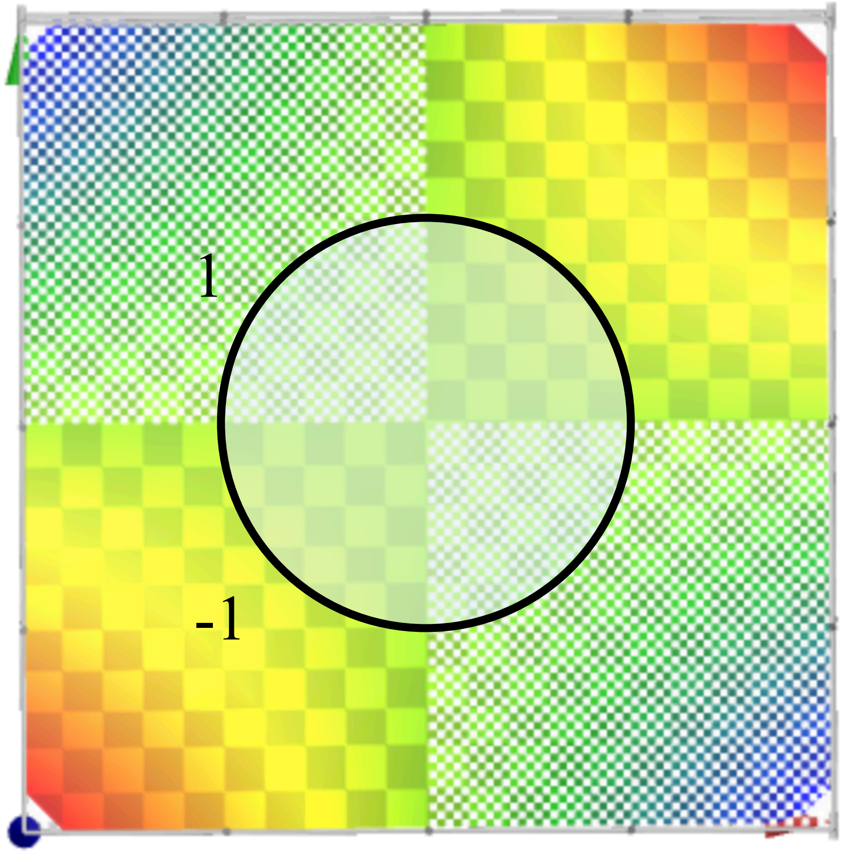
- Let X and Y be independent random variables
  - Probability Density Function (PDF) of X + Y, analogous:

$$f_{X+Y}(a) = \int_{y=-\infty}^{\infty} f_X(a-y)\, f_Y(y)\, dy$$

  - In discrete case, replace $\int_{y=-\infty}^{\infty}$ with $\sum_{y}$, and $f(y)$ with $p(y)$

Integration with Constraint

$$\iint\limits_{x^2+y^2<1} f_{x,y}\ dy\ dx = \int\limits_{-1}^{1} \int\limits_{-\sqrt{1-y^2}}^{\sqrt{1-y^2}} f_{x,y}\ dy\ dx$$



Dance, Dance Convolution

- Let X and Y be independent random variables
  - Cumulative Distribution Function (CDF) of X + Y:

$$F_{X+Y}(a) = P(X+Y \leq a)$$

$$= \iint_{x+y\leq a} f_X(x) f_Y(y)\, dx\, dy = \int_{y=-\infty}^{\infty} \int_{x=-\infty}^{a-y} f_X(x)\, dx\, f_Y(y)\, dy$$

$$= \int_{y=-\infty}^{\infty} F_X(a-y)\, f_Y(y)\, dy$$

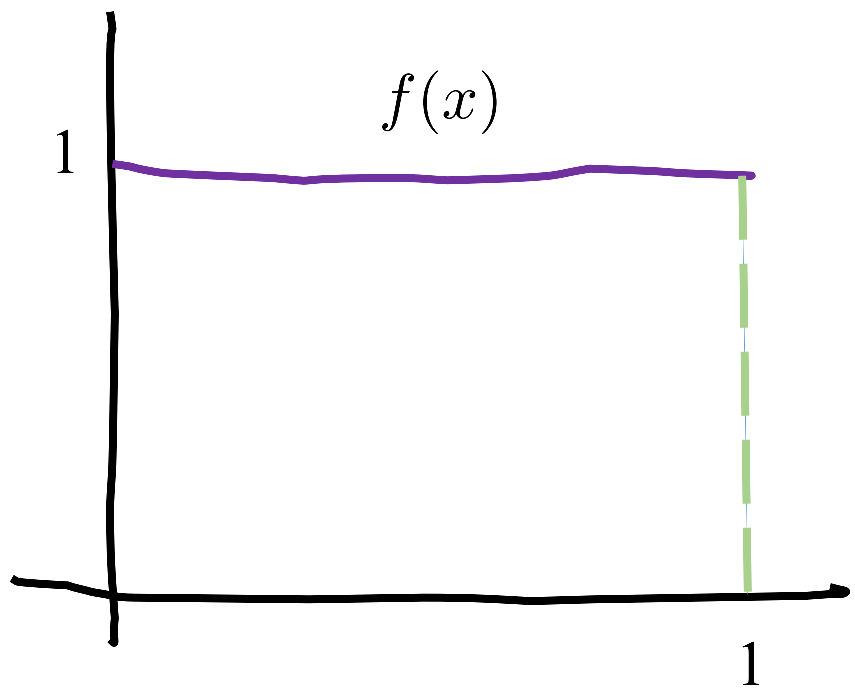
CDF of X

PDF of Y

  - In discrete case, replace $\int_{y=-\infty}^{\infty}$ with $\sum_y$, and $f(y)$ with $p(y)$

Sum of Independent Uniforms

- Let X and Y be independent random variables
  - X ~ Uni(0, 1) and Y ~ Uni(0, 1) → $f(x) = 1$ for $0 \leq x \leq 1$



For both X and Y

Sum of Independent Uniforms

- Let X and Y be independent random variables
  - X ~ Uni(0, 1) and Y ~ Uni(0, 1) → $f(x) = 1$ for $0 \le x \le 1$
  - What is PDF of X + Y?

$$f_{X+Y}(a) = \int_{y=0}^{1} f_X(a-y)\, f_Y(y)\, dy = \int_{y=0}^{1} f_X(a-y)\, dy$$

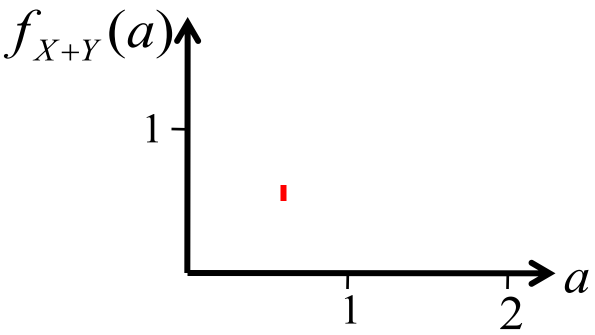
When $a = 0.5$:

$$f_{X+Y}(0.5) = \int_{y=?}^{y=?} f_X(0.5-y)dy$$

$$= \int_{0}^{0.5} f_X(0.5-y)dy$$

$$= \int_{0}^{0.5} 1dy$$

$$= 0.5$$



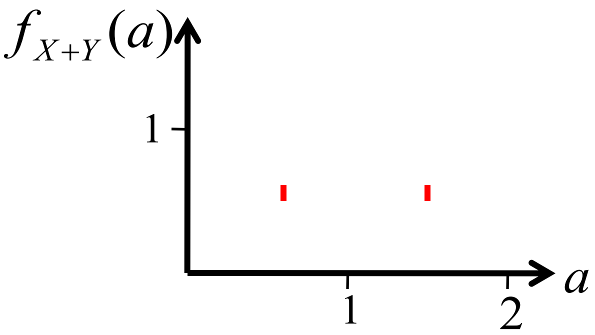
Sum of Independent Uniforms

- Let X and Y be independent random variables
  - X ~ Uni(0, 1) and Y ~ Uni(0, 1) → $f(x) = 1$ for $0 \le x \le 1$
  - What is PDF of X + Y?

$$f_{X+Y}(a) = \int_{y=0}^{1} f_X(a-y)\, f_Y(y)\, dy = \int_{y=0}^{1} f_X(a-y)\, dy$$

When $a = 1.5$:

$$\begin{aligned} f_{X+Y}(1.5) &= \int_{y=?}^{y=?} f_X(1.5-y)dy \\ &= \int_{0.5}^{1} f_X(1.5-y)dy \\ &= \int_{0.5}^{1} 1dy \\ &= 0.5 \end{aligned}$$



Sum of Independent Uniforms

- Let X and Y be independent random variables
  - X ~ Uni(0, 1) and Y ~ Uni(0, 1) → $f(x) = 1$ for $0 \leq x \leq 1$
  - What is PDF of X + Y?

$$f_{X+Y}(a) = \int_{y=0}^{1} f_X(a-y)\, f_Y(y)\, dy = \int_{y=0}^{1} f_X(a-y)\, dy$$

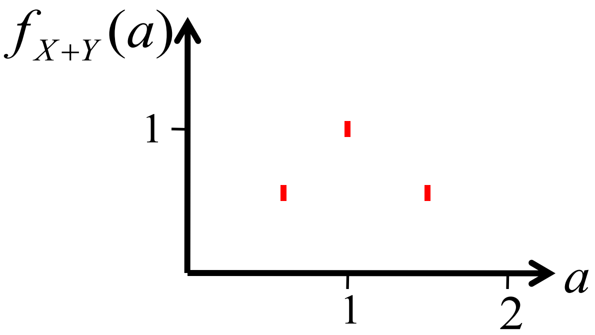
When $a = 1$:

$$f_{X+Y}(1) = \int_{y=?}^{y=?} f_X(1-y)dy$$

$$= \int_0^1 f_X(1-y)dy$$

$$= \int_0^1 1dy$$

$$= 1$$



Sum of Independent Uniforms

- Let X and Y be independent random variables
  - X ~ Uni(0, 1) and Y ~ Uni(0, 1) → $f(x) = 1$ for $0 \le x \le 1$
  - What is PDF of X + Y?

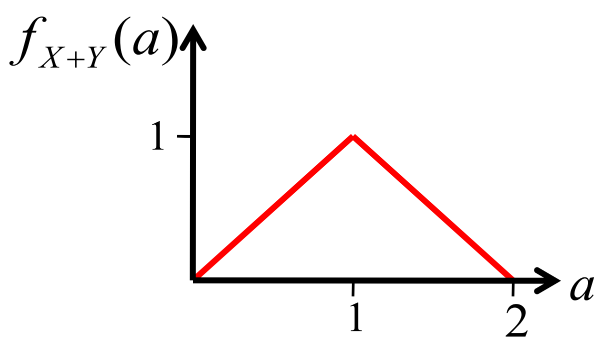
$$f_{X+Y}(a) = \int_{y=0}^{1} f_X(a-y)\, f_Y(y)\, dy = \int_{y=0}^{1} f_X(a-y)\, dy$$

  - When $0 \le a \le 1$ and $0 \le y \le a$, $0 \le a{-}y \le 1$ → $f_X(a-y) = 1$

$$f_{X+Y}(a) = \int_{y=0}^{a} dy = a$$

  - When $1 \le a \le 2$ and $a{-}1 \le y \le 1$, $0 \le a{-}y \le 1$ → $f_X(a-y) = 1$

$$f_{X+Y}(a) = \int_{y=a-1}^{1} dy = 2-a$$



  - Combining: $f_{X+Y}(a) = \begin{cases} a & 0 \le a \le 1 \\ 2-a & 1 < a \le 2 \\ 0 & \text{otherwise} \end{cases}$

Sum of Independent Normals

- Let X and Y be independent random variables
  - $X \sim N(\mu_1, \sigma_1^2)$ and $Y \sim N(\mu_2, \sigma_2^2)$
  - $X + Y \sim N(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$

- Generally, have *n* independent random variables $X_i \sim N(\mu_i, \sigma_i^2)$ for $i = 1, 2, ..., n$:

$$\left(\sum_{i=1}^{n} X_i\right) \sim N\left(\sum_{i=1}^{n} \mu_i, \sum_{i=1}^{n} \sigma_i^2\right)$$

Virus Infections

- Say you are working with the WHO to plan a response to a the initial conditions of a virus:
  - Two exposed groups
  - P1: 50 people, each independently infected with $p = 0.1$
  - P2: 100 people, each independently infected with $p = 0.4$
  - Question: Probability of more than 40 infections?

Sanity check: Should we use the Binomial Sum-of-RVs shortcut?

A. YES!
B. NO!
C. Other/none/more

Virus Infections

- Say you are working with the WHO to plan a response to a the initial conditions of a virus:
  - Two exposed groups
  - P1: 50 people, each independently infected with $p = 0.1$
  - P2: 100 people, each independently infected with $p = 0.4$
  - A = # infected in P1 A ~ Bin(50, 0.1) ≈ X ~ N(5, 4.5)
  - B = # infected in P2 B ~ Bin(100, 0.4) ≈ Y ~ N(40, 24)
  - What is P(≥ 40 people infected)?
  - P(A + B ≥ 40) ≈ P(X + Y ≥ 39.5)
  - X + Y = W ~ N(5 + 40 = 45, 4.5 + 24 = 28.5)

$$P(W \geq 39.5) = P\left(\frac{W-45}{\sqrt{28.5}} > \frac{39.5-45}{\sqrt{28.5}}\right) = 1 - \Phi(-1.03) \approx 0.8485$$

Linear Transform

$$X \sim N(\mu, \sigma^2)$$

$$Y = X + X = 2 \cdot X$$

$$Y \sim N(2\mu, 4\sigma^2)$$

$$Y = X + X = 2 \cdot X$$

$$X + X \sim N(\mu + \mu, \sigma^2 + \sigma^2)$$

$$Y \sim N(2\mu, 2\sigma^2)$$

X is not independent of X

End sum of independent vars