Math/Stat 341: Probability: Fall '21 (Williams)

Professor Steven J Miller: sjm1@williams.edu

Homepage:
https://web.williams.edu/Mathematics/sjmiller/public_html/341Fa21

Lecture 13: 10-13-21: https://youtu.be/Q-lp1yFdNvA

(2015 lecture with detailed joint PDF example: http://youtu.be/gQzorseWuVc)

Plan for the day: Lecture 13: October 13, 2021:

https://web.williams.edu/Mathematics/sjmiller/public_html/341Fa21/handouts/341Notes_Chap1.pdf

- Joint PDF
- Linearity of Expectation
- Fermat Primes
- Buffon's Needle

General items.

- Power of Linearity
- Avoiding brute force computations

Joint probability density function. Let $X_1, X_2, \ldots, X_n$ be continuous random variables with densities $f_{X_1}, f_{X_2}, \ldots, f_{X_n}$. Assume each X_i is defined on a subset of $\mathbb{R}$ (the real numbers). The joint density function of the tuple $(X_1, \ldots, X_n)$ is a non-negative, integrable function $f_{X_1,\ldots,X_n}$ such that, for every nice set $S \subset \mathbb{R}^n$ we have

$$\text{Prob}\left((X_1, \ldots, X_n) \in S\right) = \int \cdots \int_S f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) dx_1 \cdots dx_n,$$

and

$$f_{X_i}(x_i) = \int_{x_1=-\infty}^{\infty} \cdots \int_{x_{i-1}=-\infty}^{\infty} \int_{x_{i+1}=-\infty}^{\infty} \cdots \int_{x_n=-\infty}^{\infty} f_{X_1,\ldots,X_{i-1},X_{i+1},\ldots,X_n}(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \prod_{\substack{j=1 \\ j \neq i}}^{n} dx_j.$$

We call f_{X_i} the **marginal density** of X_i, and obtain it by integrating out the other $n-1$ variables.

The n random variables $X_1, \ldots, X_n$ are independent if and only if

$$f_{X_1,\ldots,X_n}(x_1, \ldots, x_n) = f_{X_1}(x_1) \cdots f_{X_n}(x_n).$$

For discrete random variables, replace the integrals with sums.

(2015 lecture with detailed joint PDF example: http://youtu.be/gQzorseWuVc)

	Prob($Y = 0$)	Prob($Y = 1$)	Prob($Y = 2$)	
Prob($X = 0$)	1/32	2/32	1/32	1/8
Prob($X = 1$)	3/32	6/32	3/32	3/8
Prob($X = 2$)	3/32	6/32	3/32	3/8
Prob($X = 3$)	1/32	2/32	1/32	1/8
	1/4	2/4	1/4	

Table 9.2: The joint density of (X, Y), where X is the number of heads in the first 3 tosses and Y is the number of heads in the last 2 tosses of 5 independent tosses of fair coins.

	Prob($V = 0$)	Prob($V = 1$)	Prob($V = 2$)	
Prob($U = 0$)	1/16	1/16	0/16	1/8
Prob($U = 1$)	2/16	3/16	2/16	3/8
Prob($U = 2$)	1/16	3/16	2/16	3/8
Prob($U = 3$)	0/16	1/16	1/16	1/8
	1/4	2/4	1/4	

Table 9.3: The joint density of (U, V), where U is the number of heads in the first 3 tosses and V is the number of heads in the last 2 tosses of 5 tosses of fair coins.

Convolutions and CDF Method

Random vars X and Y with pdf f_X and f_Y

$Z = X + Y$, assume random vars are indep, density is f_Z

Claim: $f_Z(z) = \int_{-\infty}^{\infty} f_X(t)\, f_Y(z-t)\, dt$

Want $X + Y = z$, if $X = t$ need $Y = z - t$, $\int$ over all x

$$\text{Prob}(Z \le z) = F_Z(z) = \iint_{x+y \le z} f_X(x)\, f_Y(y)\, dy\, dx$$

$$= \int_{x=-\infty}^{\infty} \int_{y=-\infty}^{z-x} f_X(x)\, f_Y(y)\, dy\, dx$$

$$= \int_{x=-\infty}^{\infty} f_X(x)\, F_Y(y) \Big|_{-\infty}^{z-x} dx = \int_{x=-\infty}^{\infty} f_X(x) F_Y(z-x)\, dx$$

$$f_Z(z) = \frac{d}{dz} F_Z(z) \overset{\text{Real Analysis}}{=} \int_{x=-\infty}^{\infty} f_X(x)\, \underbrace{f_Y(z-x)\, \frac{d}{dz}\overbrace{(z-x)}^{1}}_{\text{chain Rule}}\, dx \quad \blacksquare$$

$$(f * g)(z) = \int_{t=-\infty}^{\infty} f(t)\, g(z-t)\, dt$$

Claim $f * g = g * f$

Proof #1: change of variables

Proof #2: if f and g are densities of indep Random vars, then $f * g$ is the density of $X + Y$

However, $g * f$ is the density of $Y + X = X + Y$,

so must be the same

Theorem 9.5.1 (Linearity of Expectation) *Let $X_1, \dots, X_n$ be random variables, let $g_1, \dots, g_n$ be functions such that $\mathbb{E}[|g_i(X_i)|]$ exists and is finite, and let $a_1, \dots, a_n$ be any real numbers. Then*

$$\mathbb{E}[a_1 g_1(X_1) + \cdots + a_n g_n(X_n)] \;=\; a_1 \mathbb{E}[g_1(X_1)] + \cdots + a_n \mathbb{E}[g_n(X_n)].$$

Note the random variables are not *assumed to be independent. Also, if $g_i(X_i) = c$ (where c is a fixed number) then $\mathbb{E}[g_i(X_i)] = c$.*

If X has density f_X then $E[g(X)] = \int_{-\infty}^{\infty} g(x) f_X(x)\,dx$

$g(x) = x$ mean

$g(x) = (x-\mu)^2$ variance when $\mu = E[X]$

write $E[(X-\mu)^2] = E[(X - E[X])^2] = \sigma^2$

Study standard deviation $\sigma = \sqrt{\sigma^2}$; same units as mean

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

$$Z = X + Y$$

$$\mathbb{E}[X+Y] = \mathbb{E}[Z] = \int_{-\infty}^{\infty} z\, f_Z(z)\, dz$$

$$= \int_{z=-\infty}^{\infty} z \left[\int_{t=-\infty}^{\infty} f_X(t)\, f_Y(z-t)\, dt\right] dz$$

Assumes independence!

$$= \int_{z=-\infty}^{\infty} \int_{t=-\infty}^{\infty} (z \underbrace{-t+t}_{\text{add zero}})\, f_X(t)\, f_Y(z-t)\, dt\, dz$$

$$= \int_{t=-\infty}^{\infty} \int_{z=-\infty}^{\infty} (z-t+t)\, f_X(t)\, f_Y(z-t)\, dz\, dt$$

$$= \int_{t=-\infty}^{\infty} t \int_{z=-\infty}^{\infty} f_X(t)\, f_Y(z-t)\, dz\, dt + \int_{t=-\infty}^{\infty} f_X(t) \int_{z=-\infty}^{\infty} (z-t)\, f_Y(z-t)\, dz$$

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

If not indep:

$$E[Z] = E[X + Y] \quad \text{jpdf} \quad f_{X,Y}(x,y)$$

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

$$= \int_{x=-\infty}^{\infty} x \underbrace{\int_{y=-\infty}^{\infty} f_{X,Y}(x,y)\, dy}_{\text{marginal of } X} dx + \int_{y=-\infty}^{\infty} y \underbrace{\int_{x=-\infty}^{\infty} f_{X,Y}(x,y)\, dx}_{\text{marginal of } Y} dy$$

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

$$= E[X] + E[Y] \quad \blacksquare$$

Fermat Primes

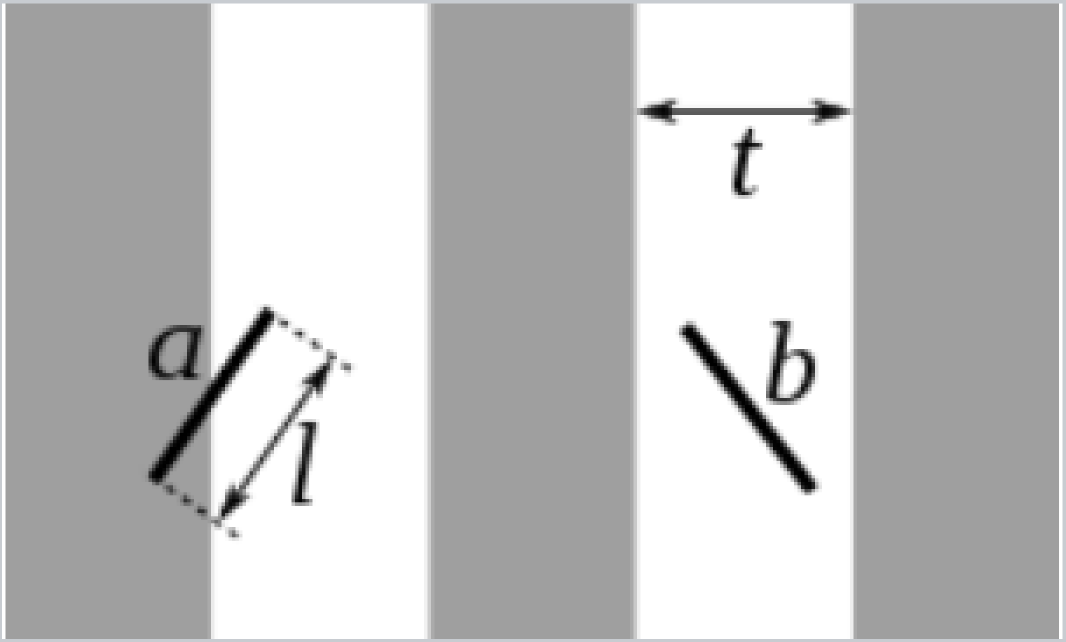
Buffon's needle problem

https://en.wikipedia.org/wiki/Buffon%27s_needle_problem

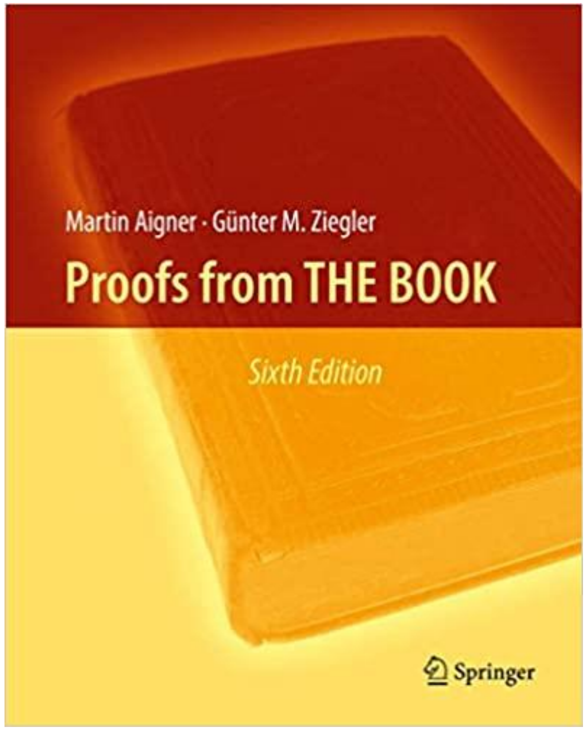
From Wikipedia, the free encyclopedia

In mathematics, **Buffon's needle problem** is a question first posed in the 18th century by Georges-Louis Leclerc, Comte de Buffon:[1]

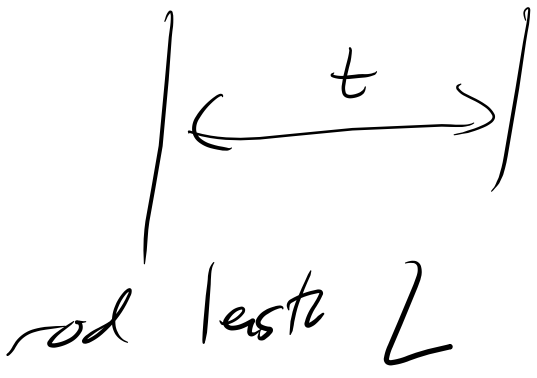
> Suppose we have a floor made of parallel strips of wood, each the same width, and we drop a needle onto the floor. What is the probability that the needle will lie across a line between two strips?



The *a* needle lies across a line, while the *b* needle does not.



https://pub.math.leidenuniv.nl/~finkelnbergh/seminarium/stelling_van_Buffon.pdf

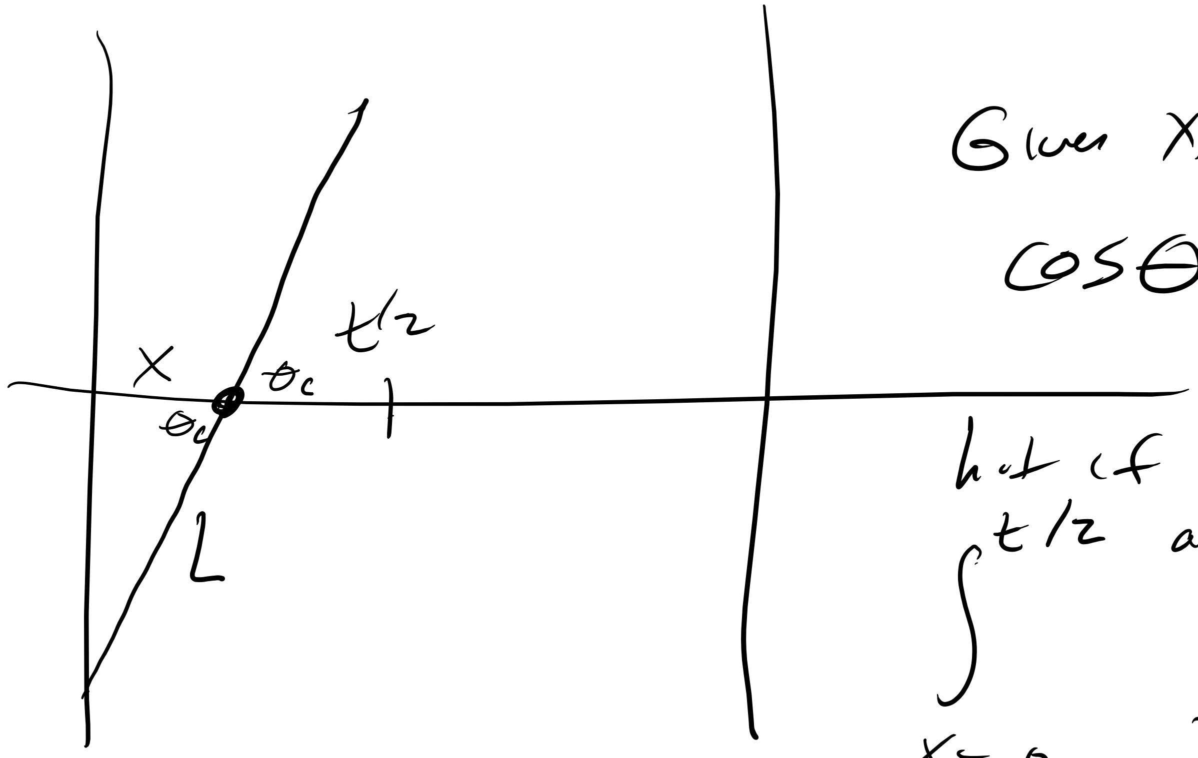


assume $L < t$; hits 0 or 1 time

rod center is at $(X, 0)$ where $0 \le X \le t/2$

angle Θ satisfies $0 \le \Theta \le \pi/2$

$$f_{X,\Theta}(x, \theta) = \frac{1}{(t/2) * (\pi/2)} = \frac{4}{t\pi}$$



Given x, L critical angle θ_c satisfies

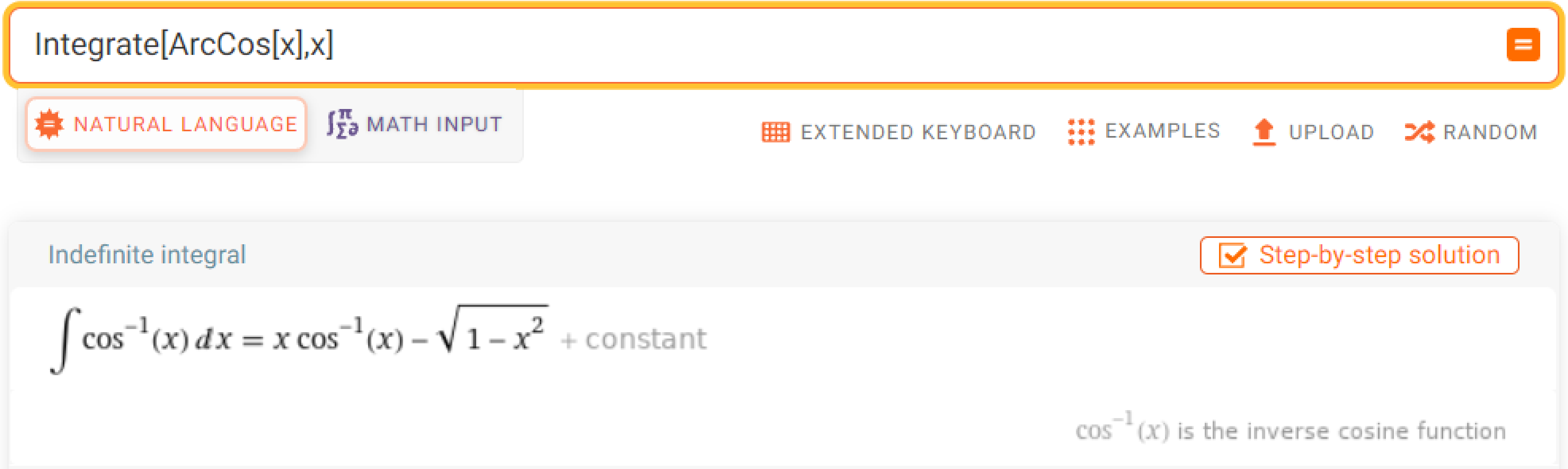
$\cos\theta_c = X/L$ so $\theta_c = \arccos(X/L)$

hit if $\theta \le \arccos(X/L)$

$$\int_{x=0}^{t/2} \int_{\theta=0}^{\arccos(x/L)} 1 \cdot \frac{4}{t\pi}\, d\theta\, dx$$

$$= \int_{x=0}^{t/2} \arccos\left(\frac{x}{L}\right) \frac{4}{t\pi}\, dx$$

WolframAlpha computational intelligence.



But maybe it is better to switch the order of integration…. What would you get?

Here is the ‘proof from the book’ link:
https://pub.math.leidenuniv.nl/~finkelnbergh/seminarium/stelling_van_Buffon.pdf