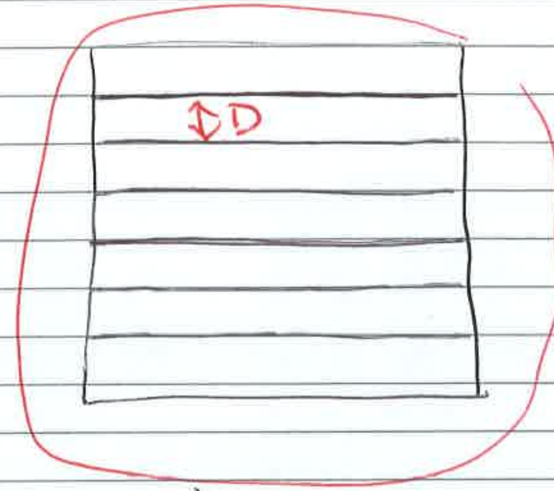
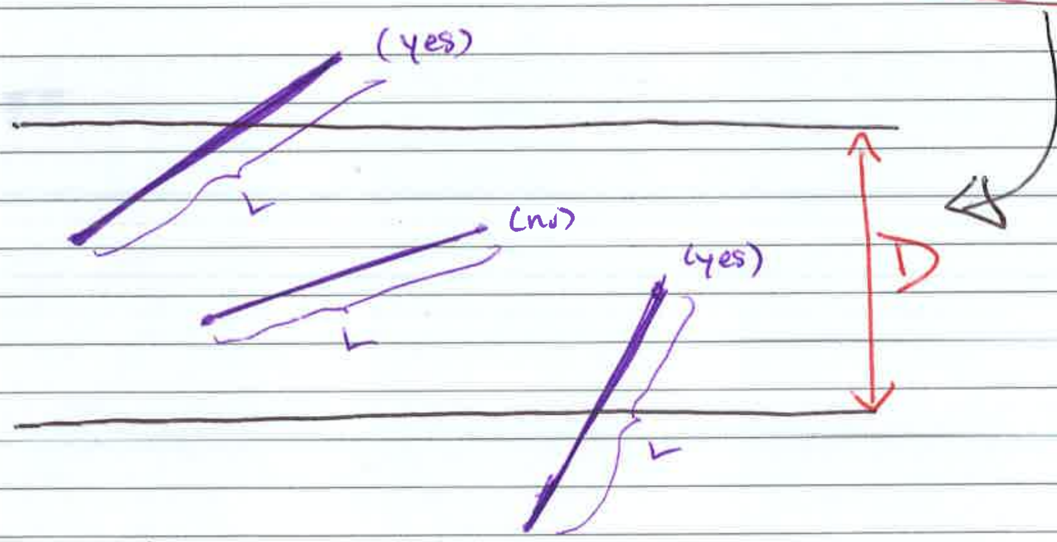
*

Buffon's Needle Experiment

Consider a [floor (hardwood) / table (ruled) / sheet of paper (lined)] with parallel lines a distance D apart.



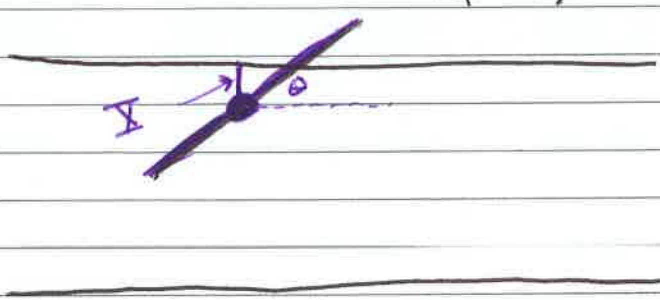
Drop a needle of length L $(L \le D)$ on the floor and ask: What is the probability that the needle will cross one of these lines?



Random Variables in this problem:

X = distance from midpoint of needle to the nearest parallel line

Θ = angle between the needle and the parallel lines.
(say top half)



Sample Space

$$S = \left\{ (x, \theta) : 0 \le x \le \frac{D}{2},\ 0 \le \theta \le \pi \right\}$$

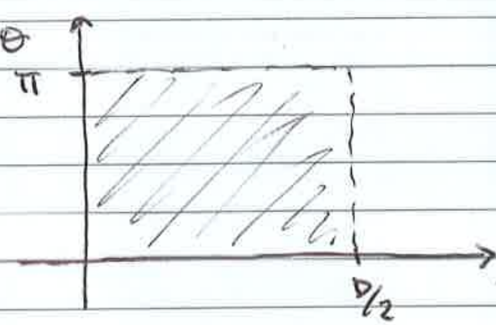
We assume that X and Θ are both uniform random variables and that they are independent. In particular

$$f_X(x) = \begin{cases} \frac{2}{D} & 0 \le x \le \frac{D}{2} \\ 0 & \text{otherwise} \end{cases}$$

$$f_\Theta(\theta) = \begin{cases} \frac{1}{\pi} & 0 \le \theta \le \pi \\ 0 & \text{otherwise} \end{cases}$$

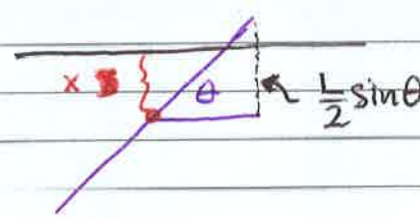
So

$$f(x,\theta) = f_X(x) \cdot f_\Theta(\theta) = \begin{cases} \frac{2}{\pi D} & 0 \le x \le \frac{D}{2},\ 0 \le \theta \le \pi \\ 0 & \text{otherwise} \end{cases}$$



A little trigonometry...

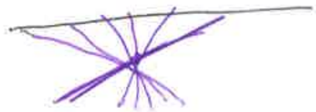
The needle crosses a line if



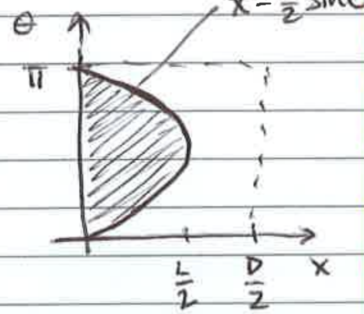
$$\frac{L}{2}\sin\theta > x$$

That is, $\frac{x}{\sin\theta} < \frac{L}{2}$

So let's determine $P\{X < \frac{L}{2}\sin\Theta\}$



$$P\left\{X < \frac{L}{2}\sin\theta\right\} = \int_0^{\pi}\int_0^{\frac{L}{2}\sin\theta} f(x,\theta)\,dx\,d\theta$$



$$= \int_0^{\pi}\int_0^{\frac{L}{2}\sin\theta} \frac{2}{\pi D}\,dx\,d\theta$$

$$= \int_0^{\pi} \frac{2}{\pi D}\cdot\frac{L}{2}\sin\theta\,d\theta$$

$$= \frac{L}{\pi D}\int_0^{\pi}\sin\theta\,d\theta$$

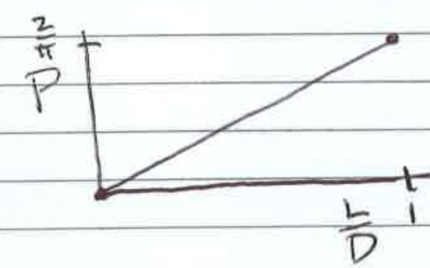
$$= \frac{L}{\pi D}\Big[-\cos\theta\Big]_0^{\pi} = \frac{L}{\pi D}\Big[-(-1)+(1)\Big] = \frac{2L}{\pi D}$$

So the probability that the needle crosses a line is

$$\boxed{P(\text{crosses line}) = \frac{2L}{\pi D}}$$

L = needle length $\leq D$
D = spacing between lines

$$\boxed{P\left(\begin{array}{c}\text{needle does not}\\ \text{cross a line}\end{array}\right) = 1 - \frac{2L}{\pi D} = \frac{\pi D - 2L}{\pi D}}$$



Further discussion of Buffon's Needle Experiment

let N_n = number of line crossings in first n needle drops.

Define the indicator variable (discrete random variable)

$$I_i = \begin{cases} 1 & \text{if needle crosses line on } i^{\text{th}} \text{ run} \\ 0 & \text{otherwise} \end{cases}$$

Then

$$N_n = \sum_{i=1}^{n} I_i$$

Note:
$$E[I_i] = 1 \cdot P\{\text{needle crosses line}\} + 0 \cdot P\{\text{needle does not cross line}\}$$

$$= \frac{2L}{\pi D} \quad \leftarrow \text{let's define this as } p$$

$$\boxed{p = \frac{2L}{\pi D}}$$

So
$$E[N_n] = E\left[\sum_{i=1}^{n} I_i\right] = \sum_{i=1}^{n} E[I_i] = \sum_{i=1}^{n} p = np$$

$$\boxed{E[N_n] = n\frac{2L}{\pi D}}$$

Also $$Var(N_n) = E[N_n^2] - (E[N_n])^2$$

side calculation: (see also text pages 158-159)

$$E[N_n^2] = E\left[\left(\sum_{i=1}^{n} I_i\right)\left(\sum_{j=1}^{n} I_j\right)\right] = E\left[\left(\sum_{i=1}^{n} I_i\right)\left(I_i + \sum_{j\neq i} I_j\right)\right]$$

$$= E\left[\sum_{i=1}^{n} I_i^2 + \sum_{i=1}^{n}\sum_{j\neq i} I_i I_j\right]$$

*

216.1

To see that this works... at least for the n=3 case...

$$\left(\sum_{i=1}^{n} I_i\right)\left(\sum_{j=1}^{n} I_j\right)$$

$$= \left(\sum_{i=1}^{n} I_i\right)\left(I_i + \sum_{j \neq i}^{n} I_j\right)$$

$$= \sum_{i=1}^{n} I_i^2 + \sum_{i=1}^{n} \sum_{j \neq i}^{n} I_i I_j$$

Maybe consider the n=3 case...

$$(I_1 + I_2 + I_3)(I_1 + I_2 + I_3)$$

$$= I_1 (I_1 + I_2 + I_3)$$

$$+ I_2 (I_2 + I_1 + I_3)$$

$$+ I_3 (I_3 + I_1 + I_2)$$

$$= \sum_{i=1}^{3} I_i^2 + \sum_{i=1}^{3} I_i \left(\sum_{j \neq i} I_j\right)$$

$$= \sum_{i=1}^{3} I_i^2 + \sum_{i=1}^{3} \sum_{j \neq i} I_i I_j$$

*

note $I_i^2 = I_i$ since I_i is an indicator variable

$$= E\left[\sum_{i=1}^{n} I_i + \sum_{i=1}^{n}\sum_{j\neq i} I_i I_j\right]$$

$$= \sum_{i=1}^{n} E[I_i] + \sum_{i=1}^{n}\sum_{j\neq i} E[I_i I_j]$$

$$= \sum_{i=1}^{n} p + \sum_{i=1}^{n}\sum_{j\neq i} E[I_i I_j]$$

$$= np + \sum_{i=1}^{n}\sum_{j\neq i} E[I_i I_j]$$

Note $I_i I_j = \begin{cases} 1 & \text{if } I_i = 1 \text{ and } I_j = 1 \\ 0 & \text{otherwise} \end{cases}$

so $E[I_i I_j] = 1 \cdot P\{I_i = 1, I_j = 1\} + 0 \cdot \text{otherwise}$

$$= P\{I_i = 1, I_j = 1\}$$

but I_i and I_j are independent so for $i \neq j$

$$P\{I_i = 1, I_j = 1\} = P\{I_i = 1\} \cdot P\{I_j = 1\} = p^2$$

Thus

$$E[N_n^2] = np + \sum_{i=1}^{n}\sum_{j\neq i}^{n}\left(p^2\right) = np + \sum_{i=1}^{n}(n-1)p^2 = np + n(n-1)p^2$$

so

$$Var(N_n) = np + n(n-1)p^2 - (np)^2 = np - np^2 = np(1-p)$$

for the needle problem

$$\boxed{Var(N_n) = n\left(\frac{2L}{\pi D}\right)\left(1 - \frac{2L}{\pi D}\right)}$$

So we may expect

$$N_n \to n\frac{2L}{\pi D} \quad \text{as } n \to \infty$$

(expected value)

$$\text{or} \quad \frac{N_n}{n} = \text{proportion of needles crossing lines} = \frac{2L}{\pi D}$$

e.g. measurable in an experiment

$$\text{So} \quad \frac{N_n D}{2nL} \text{ can be thought of as an estimator of } \frac{1}{\pi}$$

For more on this see

K.T. Siegrist : "Interactive Probability"

H. Solomon : "Geometric Probability"

An additional note on independent variables X and Y.

Proposition 2.1 :

The continuous random variables X and Y are independent if and only if their joint probability density function can be expressed as

$$f(x,y) = h(x)\,g(y) \quad \text{for all } x, y.$$

- A similar claim holds for discrete X and Y being independent iff

$$p(x,y) = h(x)\,g(y) \quad \text{for all } x, y$$

*

6.4 Conditional Distributions: Discrete Case

Def: For discrete random variables X and Y, the conditional probability mass function of X given $Y=y$ is

$$p_{X|Y}(x|y) = P\{X=x \mid Y=y\}$$

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

$$= \frac{p(x,y)}{p_Y(y)} \quad \forall y \text{ as long as } p_Y(y)>0.$$

Def: The conditional probability distribution function of X given $Y=y$ is

$$F_{X|Y}(x|y) = P\{X \le x \mid Y=y\}$$

$$= \sum_{a \le x} p_{X|Y}(a|y)$$

"sum over a with a less than x"

or $F_{X|Y}(a|y) = P\{X \le a \mid Y=y\} = \sum_{x \le a} p_{X|Y}(x|y)$ — modify notation...

- These are basically the same definitions we've used before but now everything is conditioned on $Y=y$.

If X and Y are independent, then

$$p_{X|Y}(x|y) = \frac{p_X(x)p_Y(y)}{p_Y(y)} = p_X(x)$$

*

EXAMPLE

Consider an urn with 3 red balls and 5 green balls.

2 balls chosen without replacement.

Let $X_i = \begin{cases} 1 & \text{if } i^{\text{th}} \text{ ball selected is red} \\ 0 & \text{otherwise} \end{cases}$

a). Find the joint prob. mass function for X_1, X_2.

b). Find the conditional pmf of X_2 given $X_1 = 1$.

a) $p(0,0) = \frac{5}{8} \cdot \frac{4}{7} = \frac{20}{56}$ no red (GG)

$p(0,1) = \frac{5}{8} \cdot \frac{3}{7} = \frac{15}{56}$ GR

$p(1,0) = \frac{3}{8} \frac{5}{7} = \frac{15}{56}$ RG

$p(1,1) = \frac{3}{8} \frac{2}{7} = \frac{6}{56}$ RR

$$\begin{array}{cc|cc} & & \multicolumn{2}{c}{X_2} \\ & & 0 & 1 \\ \hline X_1 & 0 & \frac{20}{56} & \frac{15}{56} \\ & 1 & \frac{15}{56} & \frac{6}{56} \end{array} \quad , \; p(x_1, x_2)$$

So $p_{X_1}(x_1) = \begin{cases} \frac{20}{56} + \frac{15}{56} = \frac{35}{56} & x_1 = 0 \\ \frac{15}{56} + \frac{6}{56} = \frac{21}{56} & x_1 = 1 \end{cases}$ ← sum over X_2 values.

marginal probability mass function for X_1

*

b) Conditional pmf of X_2 given $X_1 = 1$

$$p_{X_2|X_1}(x_2|x_1=1) = \frac{p(x_2, x_1=1)}{p_{X_1}(1)}$$

$$= \frac{p(x_2, x_1=1)}{\left(\frac{21}{56}\right)} = \begin{cases} \frac{15}{56} \Big/ \frac{21}{56} = \frac{15}{21} & x_2 = 0 \\ \frac{6/56}{\frac{21}{56}} = \frac{6}{21} & x_2 = 1 \end{cases}$$

Conditional pmf of X_2 given $X_1 = 0$

$$p_{X_2|X_1}(x_2|x_1=0) = \frac{p(x_2, x_1=0)}{p_{X_1}(0)} = \begin{cases} \frac{20}{56} \Big/ \frac{35}{56} = \frac{20}{35} & x_2 = 0 \\ \frac{\frac{15}{56}}{\frac{35}{56}} = \frac{15}{35} & x_2 = 1 \end{cases}$$