Overview

- Distribution, expected value
- LLN / Sampling
- Error in Sampling
- RNG

Randomness: Why?

What types of problems can we solve with the help of random numbers?

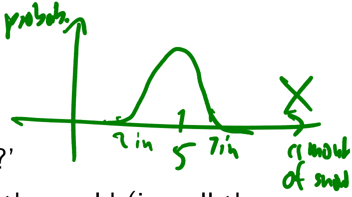
We can compute (potentially) *complicated averages*.

- Where does 'the average' web surfer end up? (PageRank)
- How much is my stock portfolio/option going to be worth?
- How will my robot behave if there is measurement error?

Random Variables

What is a random variable?

A random variable X is a function that depends on 'the (random) state of the world'.



Example: X could be

- 'how much rain tomorrow?', or
- 'will my buttered bread land face-down?'

Idea: Since I don't know the entire state of the world (i.e. all the influencing factors), I can't know the value of X.

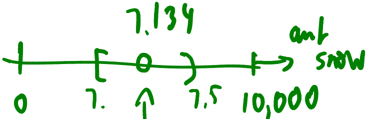
$\rightarrow$ Next best thing: Say something about the *average* case.

To do that, I need to know how likely each individual value of X is. I need a probability distribution.

Probability Distributions

What kinds of probability distributions are there?

continuous (dist/r.v.)



infinitely many possible values

each individual value has prob. 0, but ranges have finite probability;

e.g. prob$(7 < X \le 7.5) = 0.1$

discrete (dist./r.v.)

finite / discrete number of values

e.g. die 1 ... 6

p $\frac{1}{6}$ ... $\frac{1}{6}$

In both cases, probabilities must be non-neg. ≥ 0

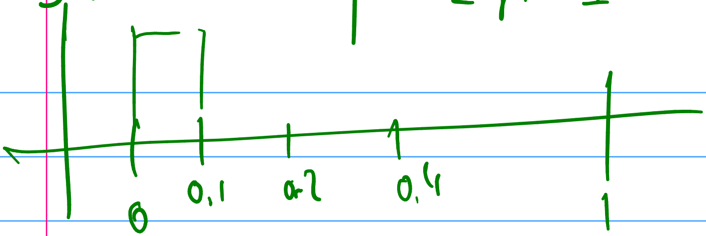
Demo: Plotting Distributions with Histograms

$p \geq 0$

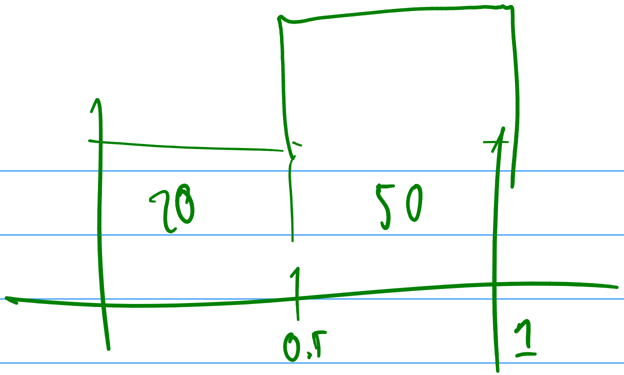
$\int p(x)\,dx = 1$

$p \geq 0$

$\sum p_i = 1$



$$P(0 \leq x \leq 0.1) + P(0.1 \leq 0.2) + P(0.2 \leq x \leq 0.4) + P(0.4 \leq x \leq 1) = 1$$


20
50
0.5
1

Expected Values/Averages: What?

Define '**expected value**' of a random variable.

discrete:

$$E[X] = \sum_{i=1}^{N} x_i \cdot p(x_i)$$

~~Define **variance** of a random variable.~~

x	1	2	3	4	5	6
$p(x)$	$1/6$	$1/6$	...			$1/6$

continuous

$$E[X] = \int x \cdot p(x) \, dx$$

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

dist. from avg. squared ← average that

Expected Value: Example I

What is the expected snowfall in Champaign?

need a time range

"a day"

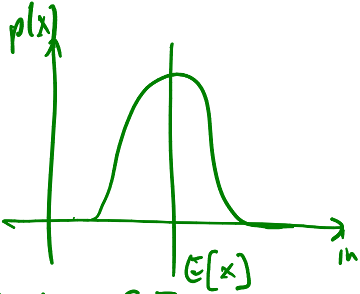
X = amount of snow

$$E[X] = \int_0^\infty X \cdot p(x)\, dx$$

amount of snow

$$\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 9 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 0.5 \\ 0 \\ 0 \end{bmatrix}$$

why approximate $E[X]$?



Tool: Law of Large Numbers

Terminology:

- *Sample*: A sample $s_1, \ldots, s_N$ of a discrete random variable X (with potential values $x_1, \ldots, x_n$) selects each s_i such that $s_i = x_j$ with probability $p(x_j)$.

In words:

- As the number of samples $N \to \infty$, the average of samples converges to the expected value with probability 1.

What can samples tell us about the distribution?

$$P\left[\lim_{N\to\infty} \boxed{\frac{1}{N}\left(\sum_{i=1}^{N} s_i\right)} = E[X]\right] = 1$$

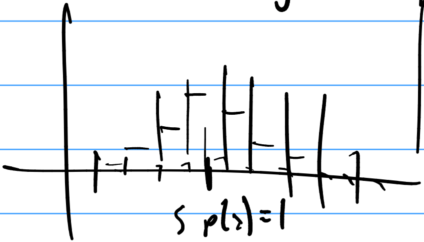
Sample mean

$$\rightarrow \quad E(X) \approx \frac{1}{N}\sum_{i=1}^{N} s_i$$

Two difficulties in applying CLN:

① - knowing the distribution

② - being able to make samples following that distribution



$\int_0^{?} p(x)\,dx = 0.7$

$\sum p(x) = 1$

Sampling: Approximating Expected Values

Integrals and sums in expected values are often challenging to evaluate.

How can we approximate an expected value?
Idea: Draw random samples. Make sure they are distributed according to $p(x)$.

What is a Monte Carlo (MC) method?

Expected Values with Hard-to-Sample Distributions

Computing the sample mean requires samples from the distribution $p(x)$ of the random variable X. What if such samples aren't available?

$$E[X] = \int x \cdot p(x)\,dx$$

$$= \int x \cdot \frac{p(x)}{\tilde{p}(x)} \cdot \tilde{p}(x)\,dx$$

$$= E\left[\hat{X} \cdot \frac{p(\hat{X})}{\tilde{p}(\hat{X})}\right]$$

Suppose $\tilde{p}(x)$ is a dist. fnct. that is easy to sample from

Let $\hat{X}$ be distributed according to $\tilde{p}(x)$.

Switching Distributions for Sampling

Found:

$$E[X] = E\left[\tilde{X} \cdot \frac{p(\tilde{X})}{\tilde{p}(\tilde{X})}\right]$$

Why is this useful for sampling?

Expected Value: Example II

What is the expected snowfall in Illinois?