Foundations of Computer Science
Lecture 18

Random Variables

Measurable Outcomes
Probability Distribution Function
Bernoulli, Uniform, Binomial and Exponential Random Variables

Last Time

1. Independence.
   - Using independence to estimate complex probabilities.
2. Coincidence.
   - FOCS-twins.
   - The birthday paradox.
   - Application to hashing.
3. Random walks and gambler's ruin.

Today: Random Variables

1. What is a random variable?
2. Probability distribution function (PDF) and Cumulative distribution function (CDF).
3. Joint probability distribution and independent random variables
4. Important random variables
   - Bernoulli: indicator random variables.
   - Uniform: simple and powerful. An equalizing force.
   - Binomial: sum of independent indicator random variables.
   - Exponential: the waiting time to the first success.

A Random Variable is a "Measurable Property"

Temperature: "measurable property" of random positions and velocities of molecules.

Toss 3 coins.

$$\text{number-of-heads(HTT)} = 1;$$
$$\text{all-tosses-match(HTT)} = 0.$$

	Sample Space Ω								
ω	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT	
$P(\omega)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	
$\mathbf{X}(\omega)$	3	2	2	1	2	1	1	0	← number of heads
$\mathbf{Y}(\omega)$	1	0	0	0	0	0	0	1	← matching tosses
$\mathbf{Z}(\omega)$	8	2	2	$\frac{1}{2}$	2	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{8}$	← H: double your money; T: halve your money

Can use random variables to define events:

$$\{\mathbf{X}=2\} = \{\text{HHT}, \text{HTH}, \text{THH}\} \qquad \mathbb{P}[\mathbf{X}=2] = \tfrac{3}{8}$$
$$\{\mathbf{X}\ge 2\} = \{\text{HHH}, \text{HHT}, \text{HTH}, \text{THH}\} \qquad \mathbb{P}[\mathbf{X}\ge 2] = \tfrac{1}{2}$$
$$\{\mathbf{Y}=1\} = \{\text{HHH}, \text{TTT}\} \qquad \mathbb{P}[\mathbf{Y}=1] = \tfrac{1}{4}$$
$$\{\mathbf{X}\ge 2 \text{ AND } \mathbf{Y}=1\} = \{\text{HHH}\} \qquad \mathbb{P}[\mathbf{X}\ge 2 \text{ AND } \mathbf{Y}=1] = \tfrac{1}{8}$$

Probability Distribution Function (PDF)

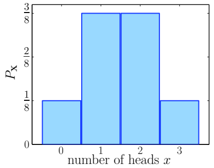
$$\underbrace{\{\text{HHH}, \text{HHT}, \text{HTH}, \text{HTT}, \text{THH}, \text{THT}, \text{TTH}, \text{TTT}\}}_{\Omega} \xrightarrow{\mathbf{X}} \underbrace{\{3, 2, 1, 0\}}_{\mathbf{X}(\Omega)}$$

Each *possible* value x of the random variable $\mathbf{X}$ corresponds to an event,

x	0	1	2	3
Event	{TTT}	{HTT, THT, TTH}	{HHT, HTH, THH}	{HHH}

For each $x \in \mathbf{X}(\Omega)$, compute $\mathbb{P}[\mathbf{X} = x]$ by adding the outcome-probabilities,

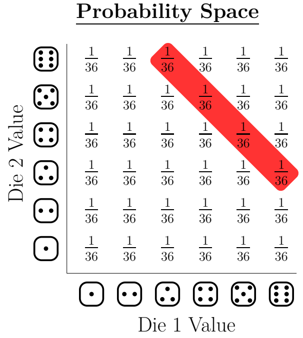
	possible values $x \in \mathbf{X}(\Omega)$			
x	0	1	2	3
$P_{\mathbf{X}}(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$



Probability Distribution Function (PDF). The probability distribution function $P_{\mathbf{X}}(x)$ is the probability for the random variable $\mathbf{X}$ to take value x,

$$P_{\mathbf{X}}(x) = \mathbb{P}[\mathbf{X} = x].$$

PDF for the Sum of Two Dice

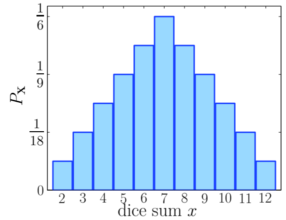


$\mathbf{X} = 9$ has four outcomes,

$$\mathbb{P}[\mathbf{X} = 9] = 4 \times \tfrac{1}{36} = \tfrac{1}{9}.$$

Possible sums are $\mathbf{X} \in \{2, 3, \ldots, 12\}$ and the PDF is

x	2	3	4	5	6	7	8	9	10	11	12
$P_{\mathbf{X}}(x)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$



Joint PDF: More Than One Random Variable

	Sample Space Ω								
ω	HHH	HHT	HTH	HTT	THH	THT	TTH	TTT	
$P(\omega)$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$	
$\mathbf{X}(\omega)$	3	2	2	1	2	1	1	0	← number of heads
$\mathbf{Y}(\omega)$	1	0	0	0	0	0	0	1	← matching tosses

$$\mathbb{P}[\mathbf{X} = 0, \mathbf{Y} = 0] = 0$$
$$\mathbb{P}[\mathbf{X} = 1, \mathbf{Y} = 0] = \tfrac{3}{8}.$$

$$P_{\mathbf{XY}}(x, y) = \mathbb{P}[\mathbf{X} = x, \mathbf{Y} = y].$$

$$\mathbb{P}[\mathbf{X} + \mathbf{Y} \le 2] = 0 + \tfrac{3}{8} + \tfrac{3}{8} + \tfrac{1}{8} + 0 = \tfrac{7}{8}.$$

$$\mathbb{P}[\mathbf{Y} = 1 \text{ AND } \mathbf{X} + \mathbf{Y} \le 2] = \tfrac{1}{8} + 0 = \tfrac{1}{8}.$$

$$\mathbb{P}[\mathbf{Y} = 1 \mid \mathbf{X} + \mathbf{Y} \le 2] = \frac{\mathbb{P}[\mathbf{Y}=1 \text{ AND } \mathbf{X}+\mathbf{Y} \le 2]}{\mathbb{P}[\mathbf{X}+\mathbf{Y} \le 2]} = \tfrac{1}{8} / \tfrac{7}{8} = \tfrac{1}{7}$$

$P_{\mathbf{XY}}(x, y)$

		$\mathbf{X}$				
		0	1	2	3	row sums
$\mathbf{Y}$	0	0	$\frac{3}{8}$	$\frac{3}{8}$	0	$\frac{3}{4}$
	1	$\frac{1}{8}$	0	0	$\frac{1}{8}$	$\frac{1}{4}$
	column sums	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

$$P_{\mathbf{Y}}(y) = \sum_{x \in \mathbf{X}(\Omega)} P_{\mathbf{XY}}(x, y)$$

$$P_{\mathbf{X}}(x) = \sum_{y \in \mathbf{Y}(\Omega)} P_{\mathbf{XY}}(x, y)$$

Independent Random Variables

Independent Random Variables measure unrelated quantities. The joint-PDF is *always* the product of the marginals.

$$P_{\mathbf{XY}}(x, y) = P_{\mathbf{X}}(x) P_{\mathbf{Y}}(y) \qquad \text{for all } (x, y) \in \mathbf{X}(\Omega) \times \mathbf{Y}(\Omega).$$

Our $\mathbf{X}$ and $\mathbf{Y}$ are *not* independent,

$P_{\mathbf{XY}}(x, y)$

		$\mathbf{X}$				
		0	1	2	3	
$\mathbf{Y}$	0	0	$\frac{3}{8}$	$\frac{3}{8}$	0	$\frac{3}{4}$
	1	$\frac{1}{8}$	0	0	$\frac{1}{8}$	$\frac{1}{4}$
		$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

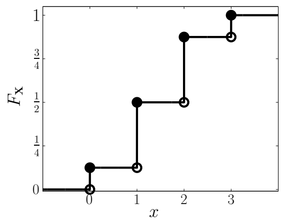
$P_{\mathbf{X}}(x) P_{\mathbf{Y}}(y)$

		$\mathbf{X}$				
		0	1	2	3	
$\mathbf{Y}$	0	$\frac{3}{32}$	$\frac{9}{32}$	$\frac{9}{32}$	$\frac{9}{32}$	$\frac{3}{4}$
	1	$\frac{1}{32}$	$\frac{3}{32}$	$\frac{3}{32}$	$\frac{1}{32}$	$\frac{1}{4}$
		$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$	

Practice: Exercise 18.4, Pop Quizzes 18.5, 18.6.

Cumulative Distribution Function (CDF)

x	0	1	2	3
$P_{\mathbf{X}}(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$
$\mathbb{P}[\mathbf{X} \le x]$	$\frac{1}{8}$	$\frac{4}{8}$	$\frac{7}{8}$	$\frac{8}{8}$



Cumulative Distribution Function (CDF). The cumulative distribution function $F_{\mathbf{X}}(x)$ is the probability for the random variable $\mathbf{X}$ to be at most x,

$$F_{\mathbf{X}}(x) = \mathbb{P}[\mathbf{X} \le x].$$

Bernoulli Random Variable: Binary Measurable (0, 1)

Two outcomes: coin toss, drunk steps left or right, etc. $\mathbf{X}$ *indicates* which outcome,

$$\mathbf{X} = \begin{cases} 1 & \text{with probability } p; \\ 0 & \text{with probability } 1 - p. \end{cases}$$

Can add Bernoullis. Toss n independent coins. $\mathbf{X}$ is the number of H.

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_n.$$

$\mathbf{X}$ is a sum of Bernoullis, each $\mathbf{X}_i$ is an independent Bernoulli.

Drunk makes n steps. Let $\mathbf{R}$ be the number of right steps

$$\mathbf{R} = \mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_n.$$

$\mathbf{R}$ is a sum of Bernoullis. $\mathbf{L} = n - \mathbf{R}$ and the final position $\mathbf{X}$ is:

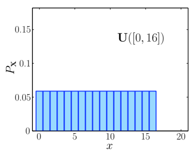
$$\mathbf{X} = \mathbf{R} - \mathbf{L} = 2\mathbf{R} - n = 2(\mathbf{X}_1 + \mathbf{X}_2 + \cdots + \mathbf{X}_n) - n.$$

Uniform Random Variable: Every Value Equally Likely

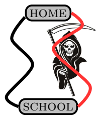
n possible values $\{1, 2, \ldots, n\}$, each with probability $\frac{1}{n}$:

$$P_{\mathbf{X}}(k) = \tfrac{1}{n}, \qquad \text{for } k = 1, \ldots, n.$$

Roll of a 6-sided *fair* die $\sim \mathbf{U}[6]$. (Uniform on $\{1, \ldots, 6\}$)



Example: Matching game (uniform is an equalizer in games of strategy).
GR will pick a path to relieve you of your lunch money.
If you pick your path uniformly, you win half the time.



Example 18.2: Guessing Larger or Smaller

I pick two numbers from $\{1, \ldots, 5\}$, as I please. I *randomly* show you one of the two, x.
You must guess if x is the larger or smaller of my two numbers.

You always say smaller: you win $\frac{1}{2}$ the time.
You say smaller if $x \le 3$ and larger if $x > 3$. I pick numbers 1,2: you win $\frac{1}{2}$ the time.

You have a strategy which wins *more* than $\frac{1}{2}$ the time, and I *cannot* prevent it!

Binomial Random Variable: Sum of Bernoullis

$\mathbf{X}$ = number of heads in n independent coin tosses with probability p of heads,

$$\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_n. \qquad \leftarrow \text{ sum of } n \text{ independent Bernoullis, } \mathbf{X}_i \sim \text{Bernoulli}(p)$$

n=5,$\mathbf{X}$=3:

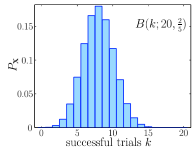
HHHTT HHTTH HTTHH TTHHH HHTHT
HTHTH THTHH HTHHT THHTH THHHT

$\leftarrow$ each has probability $p^3(1-p)^2$ (independence)

$$\mathbb{P}[\mathbf{X} = 3 \mid n = 5] = 10p^3(1-p)^5 \qquad \leftarrow \text{ add outcome probabilities}$$

In general, $\binom{n}{k}$ sequences with k heads.
Each has probability $p^k(1-p)^{n-k}$, so

$$\mathbb{P}[\mathbf{X} = k \mid n] = \binom{n}{k} p^k (1-p)^{n-k}.$$



Binomial Distribution. $\mathbf{X}$ is the number of successes in n independent trials with success probability p on each trial: $\mathbf{X} = \mathbf{X}_1 + \cdots + \mathbf{X}_n$ where $\mathbf{X}_i \sim$ Bernoulli(p).

$$P_{\mathbf{X}}(k) = B(k; n, p) = \binom{n}{k} p^k (1-p)^{n-k}.$$

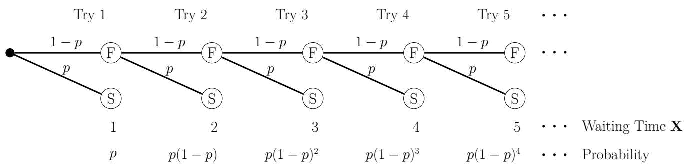
Example: guessing correctly on the multiple choice quiz: $n = 15$ questions, 5 choices ($p = \frac{1}{5}$).

number correct, k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
probability	0.035	0.132	0.231	0.250	0.188	0.103	0.043	0.014	0.003	7×10^{-4}	10^{-4}	10^{-5}	10^{-6}	~ 0	~ 0	~ 0

chances of passing are $\approx 0.4\%$

Exponential Random Variable: Waiting Time to Success

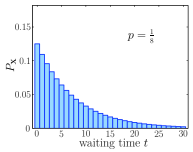
Let p be the probability to succeed on a trial.



$$\mathbb{P}[t \text{ trials}] = \mathbb{P}[\mathrm{F}^{\bullet t-1}\mathrm{S}] = (1-p)^{t-1}p$$

$$P_{\mathbf{X}}(t) = (1-p)^{t-1}p = \underbrace{\frac{p}{1-p}}_{\beta} \times (1-p)^t$$

$$= \beta(1-p)^t.$$



Example: 3 people randomly access the wireless channel. Success only if exactly one is attempting.

Try every timestep → no one succeeds. Everyone tries $\frac{1}{3}$ the time (randomly).
Success probability for *someone* is $\frac{4}{9}$. Success probability for *you* is $\frac{4}{27}$.

wait, t	1	2	3	4	5	6	7	8	9	10	11	$\cdots$
$\mathbb{P}$[someone succeeds]	0.444	0.247	0.137	0.076	0.042	0.024	0.013	0.007	0.004	0.002	0.001	$\cdots$
$\mathbb{P}$[you succeed]	0.148	0.126	0.108	0.092	0.078	0.066	0.057	0.048	0.051	0.035	0.030	$\cdots$