VE401, Probabilistic Methods in Eng.
Recitation Class - Week 2

Zhanpeng Zhou

UMJI-SJTU Joint Institute

March 4, 2021

Table of contents

Sample Space and Events

Definitions.

- **Sample space**: a space containing all possible outcomes of an experiment.
- **Event**: a subset of sample space, containing possible outcomes of the experiment.
- σ-**field**: $\mathcal{F}$ on S is a family of subsets of S s.t.
  1. $\emptyset \in \mathcal{F}$;
  2. if $A \in \mathcal{F}$, then $S \setminus A \in \mathcal{F}$;
  3. if $A_1, A_2, \ldots \in \mathcal{F}$ is a finite or countable sequence of subsets, then the union $\bigcup_k A_k \in \mathcal{F}$.

Sample Space and Events

Example 1. Suppose a coin is tossed three times. Then the sample space S contains the following possible outcomes:

$$s_1 = HHH, \quad s_2 = THH, \quad s_3 = HTH, s_4 = HHT,$$
$$s_5 = HTT, \quad s_6 = THT, \quad s_7 = TTH, s_8 = TTT,$$

where H denotes head and T denotes tail. Then the event that at most two tails are obtained is given by

Sample Space and Events

Example 1. Suppose a coin is tossed three times. Then the sample space S contains the following possible outcomes:

$$s_1 = HHH, \quad s_2 = THH, \quad s_3 = HTH, s_4 = HHT,$$
$$s_5 = HTT, \quad s_6 = THT, \quad s_7 = TTH, s_8 = TTT,$$

where H denotes head and T denotes tail. Then the event that at most two tails are obtained is given by

$$A = \{s_1, s_2, s_3, s_4, s_5, s_6, s_7\}.$$

Probability Measures and Spaces

Definition. Let S be a sample space and $\mathcal{F}$ a σ-filed on S. Then function

$$P : \mathcal{F} \to [0,1], \qquad A \mapsto P[A]$$

is a ***probability measure*** on S if

(i) $P[S] = 1$,

(ii) For any set of events $\{A_k\} \in \mathcal{F}$ such that $A_j \cap A_k = \emptyset$ for $j \neq k$,

$$P\left[\bigcup_k A_k\right] = \sum_k P[A_k].$$

The triple $(S, \mathcal{F}, P)$ is called a ***probability space***.

Probability Measures and Spaces

Properties.

- $P[S] = 1$,
- $P[\emptyset] = 0$,
- $P[S \setminus A] = 1 - P[A]$,
- $P[A_1 \cup A_2] = P[A_1] + P[A_2] - P[A_1 \cap A_2]$,

where $A, A_1, A_2, \in S$ are any events.

Cardano's Principle

Intuition: The probability of an outcome A is the proportion of the ways that can lead to A, given that all the ways are equally likely and mutually exclusive.

Cardano's Principle:

$$P[A] = \frac{\text{Number of ways leading to outcome} A}{\text{Number of ways experiments can proceed}}$$

Example 2. Tossing two (unbiased) coins, $P[\text{getting one head}]$

Cardano's Principle

Intuition: The probability of an outcome A is the proportion of the ways that can lead to A, given that all the ways are equally likely and mutually exclusive.

Cardano's Principle:

$$P[A] = \frac{\text{Number of ways leading to outcome} A}{\text{Number of ways experiments can proceed}}$$

Example 2. Tossing two (unbiased) coins, $P[\text{getting one head}] = \frac{2}{4} = 0.5$.

Basic Principles of Counting

Suppose a set A of n objects is given.

- **Permutation of k objects**: $\dfrac{n!}{(n-k)!}$ ways of choosing an <u>ordered</u> tuple of k objects from A.

- **Combination of k objects**: $\dfrac{n!}{k!(n-k)!}$ ways of choosing an <u>unordered</u> set of k objects from A.

- **Permutation of k indistinguishable objects**: $\dfrac{n!}{n_1!n_2!\ldots n_k!}$ ways of partitioning A into k disjoint subsets $A_1,\ldots,A_k$ whose union is A, where each A_i has n_i elements.

Basic Principles of Counting

Example 3. If the letters *s, s, s, t, t, t, i, i, a, c* are arranged in a random order, what is the probability that they will spell the word "statistics"?

Basic Principles of Counting

Example 3. If the letters *s, s, s, t, t, t, i, i, a, c* are arranged in a random order, what is the probability that they will spell the word "statistics"?

Solution. The problem is equivalent to arranging 5 disjoint subsets, giving the probability

$$p = 1 \left/ \frac{10!}{3! \times 3! \times 2! \times 1 \times 1} = \frac{1}{50400} \right.$$

Conditional Probability

Definitions

- **Conditional probability** of "B occurs given A has occurred":

$$P[B|A] = \frac{P[B \cap A]}{P[A]}$$

- **Total probability** for $P[B]$ on a sample space S, given events $A_1, \ldots, A_n \in S$ are mutually exclusive and $A_1 \cup \cdots \cup A_n = S$:

$$P[B] = \sum_{k=1}^{n} P[B|A_k] \cdot P[A_k].$$

- **Independence** of events A and B: $P[A \cap B] = P[A]P[B]$, which is equivalent to

$$P[A|B] = P[A] \qquad \text{if } P[B] \neq 0,$$
$$P[B|A] = P[B] \qquad \text{if } P[A] \neq 0.$$

Conditional Probability

Example 4. Suppose Keven plays a game where his score must be 1, 2, . . . , 50, and each of these 50 scores is equally likely. Suppose he get X for his first try. He then continue to play the game until he obtains another score Y such that $Y \geq X$. What is the probability of the event A that $Y = 50$?

Conditional Probability

Example 4. Suppose Keven plays a game where his score must be 1, 2, ..., 50, and each of these 50 scores is equally likely. Suppose he get X for his first try. He then continue to play the game until he obtains another score Y such that $Y \geq X$. What is the probability of the event A that $Y = 50$?

Solution. For each $i = 1, \ldots, 50$, let B_i be the event that $X = i$. Then conditional on B_i, the value of Y is equally likely to be any one of the numbers $i, i+1, \ldots, 50$. Therefore, the probability is given by

$$P[A] = \sum_{i=1}^{50} P[B_i] \cdot P[A|B_i] = \sum_{i=1}^{50} \frac{1}{50} \cdot \frac{1}{51-i} \approx 0.09.$$

Independence

Example 5. Given events A and B, what happens to $P[A \cap B]$, $P[A \cup B]$, $P[A|B]$, $P[A \neg B]$, and $P[\neg A|B]$? You are encouraged to fill out this table by yourself first.

A and B are ...	mutually exclusive	independent
$P[A \cap B]$		
$P[A \cup B]$		
$P[A\|B]$		
$P[A\|\neg B]$		
$P[\neg A\|B]$		

Independence

Example 5. Given events A and B, what happens to $P[A \cap B]$, $P[A \cup B]$, $P[A|B]$, $P[A\ \neg B]$, and $P[\neg A|B]$? You are encouraged to fill out this table by yourself first.

A and B are ...	mutually exclusive	independent
$P[A \cap B]$	0	$P[A]P[B]$
$P[A \cup B]$	$P[A] + P[B]$	$P[A] + P[B] - P[A]P[B]$
$P[A|B]$	0	$P[A]$
$P[A|\neg B]$	$\frac{P[A]}{1-P[B]}$	$P[A]$
$P[\neg A|B]$	1	$1 - P[A]$

Total Probability

Example 6. The Marriage Problem: Optimal strategy: For some $r \geq 1$, evaluate and automatically reject $r - 1$ potential partners. Then select the first candidate superior to all the previous ones, if possible.

1. Choose $r \geq 1$
2. Select k with $y_k = 1$ and $k \geq r$, if possible. Discard all others.
3. Otherwise, do not choose anyone.

The sample space can be taken to be
$S = \{(k, j) : k \text{ is selected and } x_j = 1, k = 0, \ldots, n, j = 1, \ldots, n\}$. We say that we "win" if we end up by selecting the most suitable partner, i.e., we select k with $x_k = 1$. We denote this event by

$$W_r = \{(k, k) : k = r, \ldots, n\}, \quad r \geq 1$$

Total Probability

Given $r \geq 1$, the probability of winning is denoted

$$p_r = P[W_r]$$

Goal: Choose r so that p_r is maximal.

$$p_r = P[W_r] = \sum_{m=1}^{n} P[W_r \mid B_m] P[B_m]$$

$$B_m = \{(k, m) : k = 0, \cdots, n\}, m = 1, \cdots, n.$$

Note: : $P[B_m]$: the probability that the mth person is most suitable partner "in reality".
$P[W_r|B_m]$: the probability of choosing the most suitable partner "in reality", given that the mth person is most suitable partner "in reality".

Bayes's Theorem

Theorem. Let $A_1, \ldots, A_n \subset S$ be a set of pairwise mutually exclusive events whose union is S and who each have non-zero probability of occurring. Let $B \subset S$ be any event such that $P[B] \neq 0$. Then for any $A_k, k = 1, \ldots, n$,

$$P[A_k|B] = \frac{P[B \cap A_k]}{P[B]} = \frac{P[B|A_k] \cdot P[A_k]}{\sum_{j=1}^{n} P[B|A_j] \cdot P[A_j]}.$$

Bayes's Theorem

Example 7. A box contains one fair coin and one coin with heads on both sides. Suppose one coin is selected at random and when it is tossed twice, two heads are obtained. What is the probability that the coin is the fair coin?

Bayes's Theorem

Example 7. A box contains one fair coin and one coin with heads on both sides. Suppose one coin is selected at random and when it is tossed twice, two heads are obtained. What is the probability that the coin is the fair coin?

Solution. Let E_1 be the event that the selected coin is fair, and E_2 be the event that the selected coin have two heads. Using Bayes's theorem, we have

$$
\begin{aligned}
P[E_1|HH] &= \frac{P[E_1]P[HH|E_1]}{P[HH|E_2]P[E_2] + P[HH|E_1]P[E_1]} \\
&= \frac{\dfrac{1}{2} \cdot \dfrac{1}{4}}{1 \cdot \dfrac{1}{2} + \dfrac{1}{4} \cdot \dfrac{1}{2}} = \frac{1}{5}.
\end{aligned}
$$

Random Variables and Probability Density Function

Definition. Let S be a sample space and Ω a <u>countable</u> subset of $\mathbb{R}$. A **_discrete random variable_** is a map

$$X : S \to \Omega$$

together with a function

$$f_X : \Omega \to \mathbb{R}$$

having the properties that

(i) $f_X(x) \geq 0$ for all $x \in \Omega$ and

(ii) $\displaystyle\sum_{x \in \Omega} f_X(x) = 1$.

The function f_X is called the **_probability density function_** or **_probability distribution_** of X. A random variable is given by the pair (X, f_X).

Cumulative Distribution Function

Definition. The ***cumulative distribution function*** of a random variable is defined as

$$F_X : \mathbb{R} \to \mathbb{R}, \qquad F_X(x) := P[X \leq x].$$

For a discrete random variable,

$$F_X(x) = \sum_{y \leq x} f_X(y).$$

Expectation and Variance

Definition. Let (X, f_X) be a discrete random variable.

- The **expected value** or **expectation** of X is

$$\mu_X = \mathsf{E}[X] := \sum_{x \in \Omega} x \cdot f_X(x),$$

provided that the sum (possibly series, if Ω is infinite) on the right converges absolutely.

- The **variance** is defined by

$$\sigma_X^2 = \mathsf{Var}[X] := \mathsf{E}\left[(X - \mathsf{E}[X])^2\right]$$

which is defined as long as the right-hand side exists.

- The **standard deviation** is $\sigma_X = \sqrt{\mathsf{Var}[X]}$.

Properties

- Expectation.
  - a) Suppose $\varphi : \Omega \to \mathbb{R}$ is some function, then

    $$E[\varphi \circ X] = \sum_{x \in \Omega} \varphi(x) \cdot f_X(x).$$

  - b) $E[aX + bY + c] = aE[X] + bE[Y] + c$, where $a, b, c \in \mathbb{R}$ and X, Y are random variables.
  - c) $E\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} E[X_i]$, if each expectation exists.
  - d) If $X_1, \ldots, X_n$ are independent random variables with finite expectations, and $g_i, i = 1, \ldots, n$ are functions, then

    $$E\left[\prod_{i=1}^{n} X_i\right] = \prod_{i=1}^{n} E[X_i], \quad E\left[\prod_{i=1}^{n} g_i(X_i)\right] = \prod_{i=1}^{n} E[g_i(X_i)].$$

Properties

- <u>Variance</u>.
  - **a)** $\text{Var}[X] = \text{E}[X^2] - \text{E}[X]^2$.
  - **b)** $\text{Var}[aX + b] = a^2\text{Var}[X]$, where $a, b \in \mathbb{R}$.
  - **c)** If $X_1, \ldots, X_n$ are independent random variables, then

$$\text{Var}\left[\sum_{i=1}^{n} a_i X_i\right] = \sum_{i=1}^{n} a_i^2 \text{Var}[X_i].$$

Note. If X and Y are not independent, then according to definitions,

$$\begin{aligned}
\text{Var}[X + Y] &= \text{E}\left[(X + Y - (\mu_X + \mu_Y))^2\right] \\
&= \text{E}\left[(X - \mu_X)^2\right] + \text{E}\left[(Y - \mu_Y)^2\right] + \\
&\quad + 2\text{E}\left[(X - \mu_X)(Y - \mu_Y)\right] \\
&\neq \text{Var}[X] + \text{Var}[Y].
\end{aligned}$$

Ordinary and Central Moments

Definition. The nth *(ordinary) moments* of a random variable X is given by

$$E[X^n], \qquad n \in \mathbb{N}.$$

The nth *central moments* of X is given by

$$E\left[\left(\frac{X - \mu}{\sigma}\right)^n\right], \qquad \text{where } n = 3, 4, 5, \ldots$$

Moment-Generating Function

Definition. Let (X, f_X) be a random variable and such that the sequence of moments $E[X^n], n \in \mathbb{N}$, exists. If the power series

$$m_X(t) := \sum_{k=0}^{\infty} \frac{E[X^k]}{k!} t^k$$

has radius of convergence $\varepsilon > 0$, the thereby defined function

$$m_X(t) : (-\varepsilon, \varepsilon) \to \mathbb{R}$$

is called the ***moment-generating function*** for X.

Moment-Generating Function

Theorem. Let $\varepsilon > 0$ be given such that $\mathsf{E}[e^{tX}]$ exists and has a power series expansion in t that converges for $|t| < \varepsilon$. Then the moment-generating function exists and

$$m_X(t) = \mathsf{E}[e^{tX}] \qquad \text{for } |t| < \varepsilon.$$

Furthermore,

$$E[X^k] = \left. \frac{\mathrm{d}^k m_X(t)}{\mathrm{d}t^k} \right|_{t=0}.$$

We can hence calculate the moments of X by differentiating the moment-generating function.

Moment-Generating Function

Properties.

- X is a random variable and $Y = aX + b$, $a, b \in \mathbb{R}$, then for every t such that $m_X(at)$ is finite,

$$m_Y(t) = e^{bt} m_X(at).$$

- Suppose $X_1, \ldots, X_n$ are n independent random variables, then for every value that $m_{X_i}(t)$ is finite for all $i = 1, \ldots, n$,

$$m_X(t) = \prod_{i=1}^{n} m_{X_i}(t), \qquad X = X_1 + \cdots + X_n.$$

Moment-Generating Function

Example 8. Suppose that X is a random variable with the moment-generating function

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = e^{t^2+3t}.$$

Find the mean and variance of X.

Moment-Generating Function

Example 8. Suppose that X is a random variable with the moment-generating function

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = e^{t^2+3t}.$$

Find the mean and variance of X.

Solution. We calculate

$$m_X'(t) = (2t+3)e^{t^2+3t}, \quad m_X''(t) = (2t+3)^2 e^{t^2+3t} + 2e^{t^2+3t}.$$

Therefore,

$$\mu = m_X'(0) = 3, \quad \sigma^2 = \mathsf{E}[X^2] - \mathsf{E}[X]^2 = m_X''(0) - \mu^2 = 2.$$

Bernoulli Distribution

Definition. A random variable (X, f_X) has a ***Bernoulli distribution*** with parameter $p, 0 < p < 1$ if the probability density function is defined by

$$f_X : \{0, 1\} \to \mathbb{R}, \qquad f_X(x) = \begin{cases} 1 - p, & \text{if } x = 0, \\ \\ p, & \text{if } x = 1. \end{cases}$$

Interpretation. Describe the probability of success $f_X(1)$ or failure $f_X(0)$ of a trial, given the probability of success is p.

Bernoulli Distribution

Mean, variance, and M.G.F.

- <u>Mean</u>.

$$E[X] = 0 \cdot (1 - p) + 1 \cdot p = p.$$

- <u>Variance</u>.

$$\text{Var}[X] = E[X^2] - E[X]^2 = p - p^2 = p(1 - p).$$

- <u>M.G.F.</u>

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = (1 - p) + e^t p.$$

Binomial Distribution

Definition. A random variable (X, f_X) has a **_binomial distribution_** with parameter $n \in \mathbb{N} \setminus \{0\}$ and $p, 0 < p < 1$ if it has probability density function

$$f_X : \{0, \dots, n\} \to \mathbb{R}, \qquad f_X(x) = \binom{n}{x} p^x (1-p)^{n-x}.$$

Interpretation. $f_X(x)$ is the probability of obtaining x successes in n independent and identical Bernoulli trials with parameter p.

Binomial Distribution

Mean, variance and M.G.F.

- <u>Mean</u>.

$$E[X] = \sum_{i=1}^{n} E[X_i] = np.$$

- <u>Variance</u>.

$$\text{Var}[X] = \sum_{i=1}^{n} \text{Var}[X_i] = np(1-p).$$

- <u>M.G.F.</u>

$$m_X : \mathbb{R} \to \mathbb{R}, \qquad m_X(t) = E[e^{tX}] = \prod_{i=1}^{n} E[e^{tX_i}] = (1 - p + pe^t)^n.$$

Geometric Distribution

Definition. A random variable (X, f_X) has ***geometric distribution*** with parameter $p, 0 < p < 1$ if the probability density function is given by

$$f_X : \mathbb{N} \setminus \{0\} \to \mathbb{R}, \qquad f_X(x) = (1-p)^{x-1}p.$$

Interpretation. $f_X(x)$ is the probability of x failures before the first success in the Bernoulli trials, given the probability of success for each trial is p.

Geometric Distribution

Mean, variance and M.G.F.

- <u>Mean</u>.

$$\mathsf{E}[X] = \frac{1}{p}.$$

- <u>Variance</u>.

$$\mathsf{Var}[X] = \frac{1-p}{p^2}.$$

- <u>M.G.F.</u>

$$m_X : (-\infty, -\ln(1-p)) \to \mathbb{R}, \qquad m_X(t) = \frac{pe^t}{1-(1-p)e^t}.$$

Banach Matchbox Problem

Problem setup. Suppose Keven carries two matchboxes at all times: one in his left pocket and one in his right. Each time he needs a match, he is equally likely to take it from either left or right pocket. Then one time he reaches into his pocket and discovers for the first time that the box picked is empty. Assuming each matchbox originally contained n matches, what is the probability that there are exactly k matches in the other box?

Banach Matchbox Problem

Probabilistic reasoning. Keven has to have picked $2n - k$ times, plus 1 time for him to discover that one pocket is empty. Suppose for now the empty pocket is the right one, he has to pick right for n times. The probability of this case is

$$P[L = k|R = 0] = \binom{2n - k}{n} \left(\frac{1}{2}\right)^{2n-k} \times \frac{1}{2}.$$

Since the empty pocket can be either right or left, the total probability is then given by

$$P[R = k|L = 0] + P[L = k|R = 0] = \binom{2n - k}{n} \left(\frac{1}{2}\right)^{2n-k}.$$

Banach Matchbox Problem

Counting reasoning. Suppose for now that the empty pocket is the right one, then Keven has to take $2n - k + 1$ matches to notice an empty pocket. The number of ways that he picks those matches is $\binom{2n - k}{n}$.

Using the counting method, the denominator should sum up all possible ks. Considering that the empty pocket can be either right or left, the probability of interest is given by

$$\frac{2\binom{2n-k}{n}}{2\sum_{i=0}^{n}\binom{2n-i}{n}} = \frac{\binom{2n-k}{n}}{\sum_{i=0}^{n}\binom{2n-i}{n}}.$$

What has gone wrong?

Banach Matchbox Problem

Counting reasoning. Suppose for now that the empty pocket is the right one, then Keven has to take $2n - k + 1$ matches to notice an empty pocket. The number of ways that he picks those matches is $\binom{2n-k}{n}$.

Using the counting method, the denominator should sum up all possible ks. Considering that the empty pocket can be either right or left, the probability of interest is given by

$$\frac{2\binom{2n-k}{n}}{2\sum_{i=0}^{n} \binom{2n-i}{n}} = \frac{\binom{2n-k}{n}}{\sum_{i=0}^{n} \binom{2n-i}{n}}.$$

What has gone wrong?
Note. It is important that the outcomes in the counting method are equally likely.

Banach Matchbox Problem

Counting reasoning. The sample space indicated by this reasoning is

$$S = \{k : \text{there are } k \text{ matches left in the other pocket}\},$$

which is not a ***simple sample space***. Namely, the outcomes are not equally likely. Intuitively speaking, there are more likely that the other pocket is also empty ($k = 0$) than the other pocket is totally untouched ($k = n$), given that he equally likely takes a match from left/right pocket.

Two Envelopes Problem

Problem setup. Cindy and Keven are given two indistinguishable envelopes, each of which contains a positive amount of money. One envelope contains twice as much as the other. They each may select and open one envelope and keep whatever amount it contains, but upon selection, are offered the chance to take the other envelope instead. If you are Cindy, what should you do?

Two Envelopes Problem

Ex ante argument. Once the envelopes are ready to choose, the total amount of money has already been fixed to, say Z. The probability that you select the one with smaller/larger amount is $1/2$. Therefore, the expectation should always be

$$E[X] = \frac{1}{2} \cdot \frac{1}{3}Z + \frac{1}{2} \cdot \frac{2}{3}Z = \frac{1}{2}Z,$$

which is the case for both envelopes. Any problem with this argument?

Two Envelopes Problem

Ex ante argument. Once the envelopes are ready to choose, the total amount of money has already been fixed to, say Z. The probability that you select the one with smaller/larger amount is $1/2$. Therefore, the expectation should always be

$$E[X] = \frac{1}{2} \cdot \frac{1}{3}Z + \frac{1}{2} \cdot \frac{2}{3}Z = \frac{1}{2}Z,$$

which is the case for both envelopes. Any problem with this argument? This is argument assumes that the action of opening the envelope is *uninformative*. However, the problem arises after they have selected the envelopes.

Two Envelopes Problem

Switching argument. Let x be the amount in the selected envelope. Then the probability that x is the smaller/larger amount is 0.5, with the other envelope containing $2x / \frac{1}{2}x$, respectively. Thus the expectation of the money in the other envelope is

$$\frac{1}{2} \cdot 2x + \frac{1}{2} \cdot \frac{x}{2} = \frac{5}{4}x > x.$$

Since this argument is true for both Cindy and Keven, they should both swap the envelopes. How come?

Two Envelopes Problem

Bayesian point of view. As we mentioned earlier, the paradox arises when the probability that the other player has the higher/lower amount is $1/2$ regardless whether we open the envelope. Suppose Cindy gets x in her envelope, to remain the $1/2$ probability, she has to assume that the combinations $\{x, 2x\}$ and $\{x/2, x\}$ are equally likely. This can be formalized as follows.

Two Envelopes Problem

Bayesian point of view (continued). Suppose the prior probability for the smaller amount M is f, and the envelopes contain $\{M, 2M\}$. Then we need to find the conditional probability

$$P[\text{smaller amount } M = x | \text{Cindy gets } C = x]$$
$$= \frac{P[C = x | M = x] \cdot f(x)}{P[C = x | M = x] \cdot f(x) + P[C = x | M = x/2] \cdot f(x/2)}$$
$$= \frac{f(x)}{f(x) + f(x/2)}.$$

For this to be $1/2$, we require that $f(x) = f(x/2)$ for any $x \in (0, \infty)$, which is nonsense, since that requires $f(x) = 0, \forall x \in (0, \infty)$.

Two Envelopes Problem

Bayesian point of view (continued). Then if you are Cindy, what should be the strategy? The prior probability for the two events $\{x/2, x\}$ (event A_1) and $\{x, 2x\}$ (event A_2) are $f(x/2)$ and $f(x)$. Then the expectation of the amount in the other envelope should be

$$
\begin{aligned}
E &= \frac{x}{2} \cdot P[A_1 | C = x] + 2x \cdot P[A_2 | C = x] \\
&= \frac{x}{2} \cdot \frac{f(x/2)}{f(x/2) + f(x)} + 2x \cdot \frac{f(x)}{f(x/2) + f(x)} \\
&= \frac{x}{2} \cdot \frac{f(x/2) + 4f(x)}{f(x/2) + f(x)}.
\end{aligned}
$$

Let $E > x$, then Cindy should swap the envelope if $f(x) > \frac{1}{2} f(x/2)$.

Two Envelopes Problem

Bayesian point of view (continued). From yet another point of view (if you know about total expectation), to prove that the probability of the other envelope containing larger/smaller amount is no longer 0.5, we first assume that this is the case. Denoting the amounts as X and Y.

$$E_Y[Y] = E_X\left[E_{Y|X}[Y|X]\right] = E_X\left[\frac{1}{2} \cdot \frac{1}{2}x + \frac{1}{2} \cdot 2x\right] = \frac{5}{4}E_X[X],$$

$$E_X[X] = E_Y\left[E_{X|Y}[X|Y]\right] = E_Y\left[\frac{1}{2} \cdot \frac{1}{2}y + \frac{1}{2} \cdot 2y\right] = \frac{5}{4}E_Y[Y],$$

which implies $E_X[X] = E_Y[Y] = 0$.

Two Envelopes Problem

Bayesian point of view (continued). From yet another point of view (if you know about total expectation), to prove that the probability of the other envelope containing larger/smaller amount is no longer 0.5, we first assume that this is the case. Denoting the amounts as X and Y.

$$E_Y[Y] = E_X\left[E_{Y|X}[Y|X]\right] = E_X\left[\frac{1}{2} \cdot \frac{1}{2}x + \frac{1}{2} \cdot 2x\right] = \frac{5}{4}E_X[X],$$

$$E_X[X] = E_Y\left[E_{X|Y}[X|Y]\right] = E_Y\left[\frac{1}{2} \cdot \frac{1}{2}y + \frac{1}{2} \cdot 2y\right] = \frac{5}{4}E_Y[Y],$$

which implies $E_X[X] = E_Y[Y] = 0$.

Barry Nalebuff, *The Other Person's Envelope is Always Greener*.

Exercise 1.

Exercise 1. Suppose n tennis players enter a tournament. In each round of the play, if the number of players is odd, then one player is randomly selected to enter the next round. If the number of players is even, all players are randomly paired. The loser in each pair is eliminated from the tournament. This process ends until the final winner is determined. Then what is the probability that two specific players A and B will ever play against each other?

Exercise 1.

Solution. The number of players eliminated will be the same as the number of matches. Therefore, exactly $n-1$ matches must be played. The number of possible pairs of players is $\binom{n}{2}$. Each of the two players in every match is equally likely to win. Thus before the tournament begins, every possible pair of players is equally likely to appear in each particular one of the $n-1$ matches. Accordingly, the probability that players A and B will meet in some particular match that is specified in advance is $1 \Big/ \binom{n}{2}$. Since there are $n-1$ matches where they might meet, the probability is

$$p = (n-1) \Big/ \binom{n}{2} = \frac{2}{n}.$$

Exercise 2.

Exercise 2. Suppose Keven plays the game of craps as follows. He first rolls two dice, and the sum x of the two numbers is observed.

- If $x \in \{7, 11\}$, he wins immediately.
- If $x \in \{2, 3, 12\}$, he loses immediately.
- Otherwise, the dice are rolled again and again until either the sum x or 7 appear. He wins if x first appears, and loses if 7 first appears.

What is the probability of winning?

Exercise 2. 1

Solution. Denote by W the event that Keven wins the game. Then the sample space S contains all possible sequences of sums from the rolls of dice. Let B_i the event that the sum of the first roll is i, where $i = 2, \ldots, 12$. Then according to the total probability formula,

$$P(W) = \sum_{i=2}^{12} P[B_i] P[W|B_i].$$

Then we need to specify $P[B_i]$ and $P[W|B_i]$. From the game rule,

$$P[W|B_2] = P[W|B_3] = P[W|B_{12}] = 0, \qquad P[W|B_7] = P[W|B_{11}] = 1.$$

For the cases where $i \in \{4, 5, 6, 8, 9, 10\}$, the probability of winning is given by

$$P[W|B_i] = \frac{P[B_i]}{P[B_i \cup B_7]},$$

Exercise 2. II

since the game ends either by rolling a sum of *i* or 7. We can obtain

$$P[W|B_4] = \frac{3/36}{3/36 + 6/36} = \frac{1}{3}, \qquad P[W|B_5] = \frac{4/36}{4/36 + 6/36} = \frac{2}{5},$$

$$P[W|B_6] = \frac{5/36}{5/36 + 6/36} = \frac{5}{11}, \qquad P[W|B_8] = \frac{5}{36}5/36 + 6/36 = \frac{5}{11},$$

$$P[W|B_9] = \frac{4/36}{4/36 + 6/36} = \frac{2}{5}, \qquad P[W|B_{10}] = \frac{3}{36}3/36 + 6/36 = \frac{1}{3}.$$

Therefore, the total probability is given by

$$P[W] = \sum_{i=2}^{12} P[B_i]P[W|B_i] = 0.493.$$

Exercise 3.

Exercise 3 (*The Gambler's Ruin Problem*). Suppose Keven and Cindy are playing a game against each other. Let $p \in (0, 1)$ denotes the probability that Keven wins each play of the game. If Keven wins one play, he wins one dollar from Cindy, and if he loses, he loses one dollar. Suppose Keven and Cindy start from i and $k - i$ dollars, respectively. Keven has decided to quit the game as soon as his current fortune reaches either k or 0. Then what is the probability that he finally reaches k dollars?

Exercise 3. 1

Solution. Starting with i and $k - i$ dollars respectively, we denote the probability that Keven will reach k dollars as a_i. Due to independence, the setup of the game is initialized with the change only being the starting fortune. Namely, suppose after the first play, Keven owns j dollars, then we get the conditional probability

$$P[W|(i,j)] = a_j,$$

where W denotes the event that Keven ends up with k dollars, and (i, j) denotes his sequence of fortune.

Particularly, if the sequence reaches 0, then the probability is 0, giving $a_0 = 0$. Similarly, $a_k = 1$. Let A_1 denote the event that Keven wins the first play, and let B_1 denote the event that he loses the first play. Then we obtain

$$\begin{aligned} P[W] &= P[A_1]P[W|A_1] + pB_1P[W|B_1] \\ &= pP[W|A_1] + (1-p)P[W|B_1], \end{aligned}$$

Exercise 3. II

namely,

$$a_i = pa_{i+1} + (1-p)a_{i-1}, \qquad i = 1, \ldots, k-1$$

Therefore, we obtain

$$a_1 = pa_2,$$
$$a_2 = pa_3 + (1-p)a_1,$$
$$a_3 = pa_4 + (1-p)a_2,$$
$$\vdots$$
$$a_{k-2} = pa_{k-1} + (1-p)a_{k-3},$$
$$a_{k-1} = p + (1-p)a_{k-2}.$$

Since $a_i = pa_i + (1-p)a_i$, we have

$$a_2 - a_1 = \frac{1-p}{p}a_1,$$

$$a_3 - a_2 = \frac{1-p}{p}(a_2 - a_1) = \left(\frac{1-p}{p}\right)^2 a_1,$$

$$a_4 - a_3 = \frac{1-p}{p}(a_3 - a_2) = \left(\frac{1-p}{p}\right)^3 a_1,$$

$$\vdots$$

$$a_{k-1} - a_{k-2} = \frac{1-p}{p}(a_{k-2} - a_{k-3}) = \left(\frac{1-p}{p}\right)^{k-2} a_1,$$

$$1 - a_{k-1} = \frac{1-p}{p}(a_{k-1} - a_{k-2}) = \left(\frac{1-p}{p}\right)^{k-1} a_1.$$

Exercise 3. IV

Summing up all terms, we have

$$1 - a_1 = a_1 \sum_{i=1}^{k-1} \left(\frac{1-p}{p} \right)^i.$$

Then based on the value of p, we have to discuss two cases.

- $p = \dfrac{1}{2}$, (fair game). This gives

$$1 - a_1 = (k-1)a_1 \qquad \Rightarrow \qquad a_1 = \frac{1}{k},$$

which then indicates

$$a_i = \frac{i}{k}, \qquad i = 1, \ldots, k-1$$

by induction.

Exercise 3. V

- $p \neq \frac{1}{2}$, (unfair game). In this case, we have

$$
1 - a_1 = a_1 \cdot \frac{\left(\dfrac{1-p}{p}\right)^k - \left(\dfrac{1-p}{p}\right)}{\dfrac{1-p}{p} - 1} \qquad \Rightarrow \qquad a_1 = \frac{\dfrac{1-p}{p} - 1}{\left(\dfrac{1-p}{p}\right)^k - 1}.
$$

Then similarly, we have

$$
a_i = \frac{\left(\dfrac{1-p}{p}\right)^i - 1}{\left(\dfrac{1-p^k}{p}\right)^k - 1}, \qquad i = 1, \dots, k-1.
$$