S&DS 241 Lecture 4

Law of total probability. Bayes formula.

B-H: Sec 2.3,2.4,2.7,2.8

Law of total probability: Motivations

Example: sampling without replacements (Lec03)

Consider an urn consisting of $\{{\color{blue}\bullet\bullet\bullet}{\color{red}\bullet\bullet}\}$. Draw one ball at random, remove it; then draw another at random. Find:

$$P(\text{second ball is } {\color{blue}\bullet})$$

Example: sampling without replacements (Lec03)

Consider an urn consisting of $\{$●●●●●$\}$. Draw one ball at random, remove it; then draw another at random. Find:

$$P(\text{second ball is ●})$$

Thought process

- With a chance of $\frac{3}{5}$, first ball is ●, in which case there is a $\frac{1}{2}$ chance that the second ball is ●

Example: sampling without replacements (Lec03)

Consider an urn consisting of {●●●●●}. Draw one ball at random, remove it; then draw another at random. Find:

$$P(\text{second ball is } ●)$$

Thought process

- With a chance of $\frac{3}{5}$, first ball is ●, in which case there is a $\frac{1}{2}$ chance that the second ball is ●
- With a chance of $\frac{2}{5}$, first ball is ●, in which case there is a $\frac{3}{4}$ chance that the second ball is ●

Example: sampling without replacements (Lec03)

Consider an urn consisting of $\{{\color{blue}\bullet}{\color{blue}\bullet}{\color{blue}\bullet}{\color{red}\bullet}{\color{red}\bullet}\}$. Draw one ball at random, remove it; then draw another at random. Find:

$$P(\text{second ball is } {\color{blue}\bullet})$$

<u>Thought process</u>

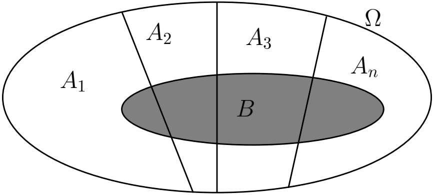
- With a chance of $\frac{3}{5}$, first ball is ${\color{blue}\bullet}$, in which case there is a ${\color{red}\frac{1}{2}}$ chance that the second ball is ${\color{blue}\bullet}$
- With a chance of $\frac{2}{5}$, first ball is ${\color{red}\bullet}$, in which case there is a ${\color{red}\frac{3}{4}}$ chance that the second ball is ${\color{blue}\bullet}$
- So overall,

$$P(\text{second ball is } {\color{blue}\bullet}) = \frac{3}{5} \times {\color{red}\frac{1}{2}} + \frac{2}{5} \times {\color{red}\frac{3}{4}} = \frac{3}{5}$$

Strategy: divide and conquer

1. List all possible scenarios exhaustively
2. Compute the conditional probability of the event under each scenario
3. Take the average (weighted by the likelihood of each scenario)

Law of total probability (LOTP)



- Let $A_1, \ldots, A_n$ be a partition of Ω, i.e.,

$$A_1 \cup \cdots \cup A_n = \Omega \quad \text{and} \quad A_i \cap A_j = \varnothing \text{ for } i \neq j$$

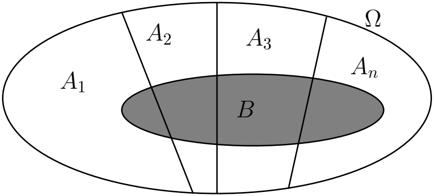
- Then

$$\boxed{P(B) = \sum_{i=1}^{n} P(B|A_i)\, P(A_i)}$$

- Special case $n = 2$: conditioned on whether A occurs or not,

$$P(B) = P(B|A)\, P(A) + P(B|A^c)\, P(A^c)$$

Proof of LOTP



Proof.

$$B = B \cap \Omega = B \cap (A_1 \cup \cdots \cup A_n) = \underbrace{(A_1 \cap B) \cup \cdots \cup (A_n \cap B)}_{\text{mutually exclusive}}$$

$$\Rightarrow P(B) = \sum_{i=1}^{n} P(A_i \cap B) = \sum_{i=1}^{n} P(B|A_i)P(A_i)$$

$\square$

Bayes formula/rule/theorem

Bayes formula/rule/theorem

Recall

- Independence is a symmetric notion
- Conditioning is an asymmetric notion:

$$P(A|B) \neq P(B|A) \quad \text{in general}$$

Bayes formula/rule/theorem

Recall

- Independence is a symmetric notion
- Conditioning is an asymmetric notion:

$$P(A|B) \neq P(B|A) \quad \text{in general}$$

- How to find $P(A|B)$ based on $P(B|A)$?

Bayes formula/rule/theorem

Recall

- Independence is a symmetric notion
- Conditioning is an asymmetric notion:

$$P(A|B) \neq P(B|A) \quad \text{in general}$$

- How to find $P(A|B)$ based on $P(B|A)$?

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}}$$

Bayes formula/rule/theorem

Recall

- Independence is a symmetric notion
- Conditioning is an asymmetric notion:

$$P(A|B) \neq P(B|A) \quad \text{in general}$$

- How to find $P(A|B)$ based on $P(B|A)$?

$$\boxed{P(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{P(B|A)P(A)}{P(B)}}$$

- Confusing these $P(A|B)$ and $P(B|A)$ is called the prosecutor's fallacy

Example (B-H Sec 2.8)

In 1998, Sally Clark was tried for murder after two of her sons died shortly after birth. During the trial, an expert witness for the prosecution testified that the probability of a newborn dying of sudden infant death syndrome (SIDS) was $1/8500$, so the probability of two deaths due to SIDS in one family was $(1/8500)^2$, or about 1/73 million. Therefore, he continued, the probability of Clark's innocence was 1/73 million.

Example (B-H Sec 2.8)

In 1998, Sally Clark was tried for murder after two of her sons died shortly after birth. During the trial, an expert witness for the prosecution testified that the probability of a newborn dying of sudden infant death syndrome (SIDS) was $1/8500$, so the probability of two deaths due to SIDS in one family was $(1/8500)^2$, or about 1/73 million. Therefore, he continued, the probability of Clark's innocence was 1/73 million.

Two issues:

- Confusing $P(\text{evidence}|\text{innocence})$ with $P(\text{innocence}|\text{evidence})$:

$$P(\text{innocence}|\text{evidence}) = \frac{P(\text{evidence}|\text{innocence})P(\text{innocence})}{P(\text{evidence})}$$

Example (B-H Sec 2.8)

In 1998, Sally Clark was tried for murder after two of her sons died shortly after birth. During the trial, an expert witness for the prosecution testified that the probability of a newborn dying of sudden infant death syndrome (SIDS) was $1/8500$, so the probability of two deaths due to SIDS in one family was $(1/8500)^2$, or about 1/73 million. Therefore, he continued, the probability of Clark's innocence was 1/73 million.

Two issues:

- Confusing $P(\text{evidence}|\text{innocence})$ with $P(\text{innocence}|\text{evidence})$:

$$P(\text{innocence}|\text{evidence}) = \frac{P(\text{evidence}|\text{innocence})P(\text{innocence})}{P(\text{evidence}|\text{innocence})P(\text{innocence}) + P(\text{evidence}|\text{guilty})P(\text{guilty})}$$

Example (B-H Sec 2.8)

In 1998, Sally Clark was tried for murder after two of her sons died shortly after birth. During the trial, an expert witness for the prosecution testified that the probability of a newborn dying of sudden infant death syndrome (SIDS) was $1/8500$, so the probability of two deaths due to SIDS in one family was $(1/8500)^2$, or about 1/73 million. Therefore, he continued, the probability of Clark's innocence was 1/73 million.

Two issues:

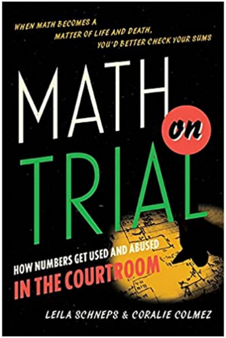
- Confusing $P(\text{evidence}|\text{innocence})$ with $P(\text{innocence}|\text{evidence})$:

$$P(\text{innocence}|\text{evidence}) = \frac{P(\text{evidence}|\text{innocence})P(\text{innocence})}{P(\text{evidence}|\text{innocence})P(\text{innocence}) + P(\text{evidence}|\text{guilty})P(\text{guilty})}$$

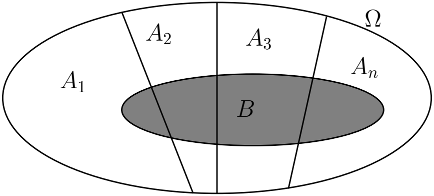
- $P(\text{evidence}|\text{innocence}) = \frac{1}{8500^2}$ assumes independence (questionable).

More on this

- *People v. Collins* in Grinstead-Snell, Section 4.1, Problem 28.
- For more in-depth discussion, see https://en.wikipedia.org/wiki/Prosecutor's_fallacy and



Summary

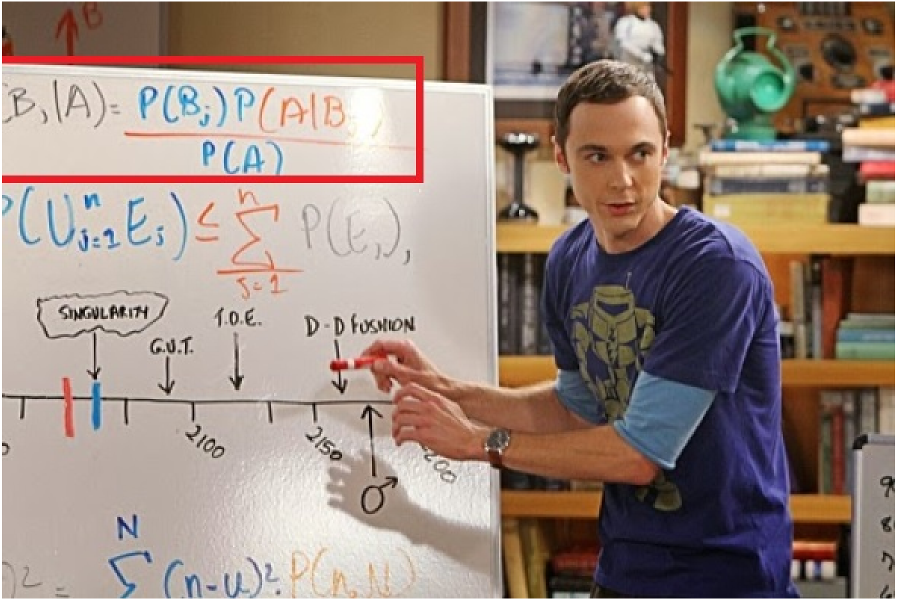


- Law of total probability:

$$P(B) = \sum_{i=1}^{n} P(B|A_i)\, P(A_i)$$

- Bayes formula:

$$P(A_j|B) = \frac{P(B|A_j)\, P(A_j)}{\sum\limits_{i=1}^{n} P(B|A_i)\, P(A_i)}$$


SINGULARITY
G.U.T.
T.O.E.
D-D FUSION
2100
2150
P(A)

Statistical inference: hypotheses testing

- $A_1, \ldots, A_n$: hypotheses
- B: result of the experiment, observed data
- $P(A_i)$: prior belief of the ith hypothesis
- Use Bayes formula to compute the posterior probability

$$P(j\text{th hypothesis is true}|\text{data}) = P(A_j|B) = \frac{P(B|A_j)\,P(A_j)}{\sum_{i=1}^{n} P(B|A_i)\,P(A_i)}$$

Rare disease

A doctor gives a patient a test for a particular cancer. Before the test, the only evidence the doctor has to go on is that 1 person in 1000 has this cancer. Experience has shown that, in 99% of the cases in which cancer is present, the test is positive; and in 95% of the cases in which cancer is not present, it is negative. If the test turns out to be positive, what probability should the doctor assign to the event that cancer is present?

Mathematical description

Let $+$ denote test positive and $-$ test negative.

$$
\begin{aligned}
P\left(\text{cancer}\right) &= 1/1000, \quad P\left(\text{no cancer}\right) = 999/1000, \\
\underbrace{P\left(+|\text{cancer}\right)}_{\text{true positive}} &= 99\%, \quad \underbrace{P\left(-|\text{cancer}\right)}_{\text{false negative}} = 1\% \\
\underbrace{P(-|\text{no cancer})}_{\text{true negative}} &= 95\%, \quad \underbrace{P(+|\text{no cancer})}_{\text{false positive}} = 5\%
\end{aligned}
$$

Mathematical description

Let $+$ denote test positive and $-$ test negative.

$$
\begin{aligned}
P\,(\text{cancer}) &= 1/1000, \quad P\,(\text{no cancer}) = 999/1000, \\
\underbrace{P\,(+|\text{cancer})}_{\text{true positive}} &= 99\%, \quad \underbrace{P\,(-|\text{cancer})}_{\text{false negative}} = 1\% \\
\underbrace{P(-|\text{no cancer})}_{\text{true negative}} &= 95\%, \quad \underbrace{P(+|\text{no cancer})}_{\text{false positive}} = 5\%
\end{aligned}
$$

The question is

$$P\,(\text{cancer}|+) = ?$$

Mathematical description

Let $+$ denote test positive and $-$ test negative.

$$\begin{aligned}
P\left(\text{cancer}\right) &= 1/1000, \quad P\left(\text{no cancer}\right) = 999/1000, \\
\underbrace{P\left(+|\text{cancer}\right)}_{\text{true positive}} &= 99\%, \quad \underbrace{P\left(-|\text{cancer}\right)}_{\text{false negative}} = 1\% \\
\underbrace{P(-|\text{no cancer})}_{\text{true negative}} &= 95\%, \quad \underbrace{P(+|\text{no cancer})}_{\text{false positive}} = 5\%
\end{aligned}$$

The question is

$$P\left(\text{cancer}|+\right) = ?$$

Grinstead-Snell:

> *When a group of second-year medical students was asked this question, over half of the students incorrectly guessed the probability to be greater than 50%*

Compute the posterior probability

$$
\begin{aligned}
& P(\text{cancer}|+) \\
= \quad & \frac{P(\text{cancer}, +)}{P(+)} \\
= \quad & \frac{P(\text{cancer}, +)}{P(\text{cancer}, +) + P(\text{no cancer}, +)} \\
= \quad & \frac{P(\text{cancer})\, P(+|\text{cancer})}{P(\text{cancer})\, P(+|\text{cancer}) + P(\text{no cancer})\, P(+|\text{no cancer})} \\
= \quad & \frac{1/1000 \times 99\%}{1/1000 \times 99\% + 999/1000 \times 5\%} \\
= \quad & 0.0194
\end{aligned}
$$

Compute the posterior probability

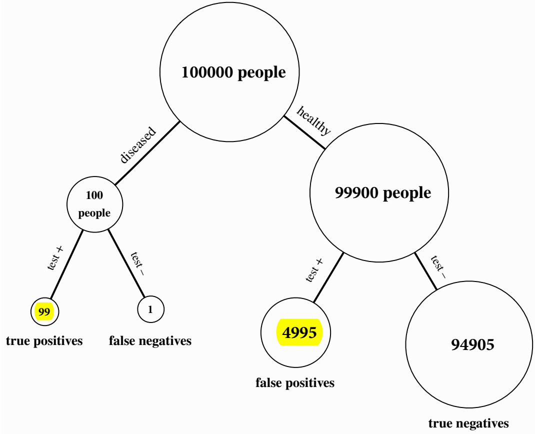
$$
\begin{aligned}
& P\left(\text{cancer}|+\right) \\
= \ & \frac{P\left(\text{cancer},+\right)}{P\left(+\right)} \\
= \ & \frac{P\left(\text{cancer},+\right)}{P\left(\text{cancer},+\right)+P\left(\text{no cancer},+\right)} \\
= \ & \frac{P\left(\text{cancer}\right)P\left(+|\text{cancer}\right)}{P\left(\text{cancer}\right)P\left(+|\text{cancer}\right)+P\left(\text{no cancer}\right)P\left(+|\text{no cancer}\right)} \\
= \ & \frac{1/1000\times 99\%}{1/1000\times 99\%+999/1000\times 5\%} \\
= \ & 0.0194
\end{aligned}
$$

Are you surprised?

- The test seems very reliable, but why the probability 1.94% is so small?
- Given the probability is 1.94%, should the patient not be alarmed?

Be careful when priors are very biased

- Tree-diagram (see B-H p. 57):



- In fact the test result has increased the likelihood almost 20-fold! (from 0.001 to 0.0194)

More reading on this example

- Blitzstein-Hwang: Example 2.3.9 (pp. 56-58)

Example: Deuce

Alice plays tennis against Bob. The game is at deuce. Suppose

- Alice wins each point with probability p and loses with probability $q = 1 - p$
- Each point is played independently
- The game is won by the player who leads by 2 points

What is the probability that Alice eventually wins the game?

Example: Deuce

Alice plays tennis against Bob. The game is at deuce. Suppose

- Alice wins each point with probability p and loses with probability $q = 1 - p$
- Each point is played independently
- The game is won by the player who leads by 2 points

What is the probability that Alice eventually wins the game?

For example

- WLWW $\rightarrow$ Alice wins

Example: Deuce

Alice plays tennis against Bob. The game is at deuce. Suppose

- Alice wins each point with probability p and loses with probability $q = 1 - p$
- Each point is played independently
- The game is won by the player who leads by 2 points

What is the probability that Alice eventually wins the game?

For example

- WLWW $\rightarrow$ Alice wins
- WLLWLL $\rightarrow$ Alice loses

Example: Deuce

Alice plays tennis against Bob. The game is at deuce. Suppose

- Alice wins each point with probability p and loses with probability $q = 1 - p$
- Each point is played independently
- The game is won by the player who leads by 2 points

What is the probability that Alice eventually wins the game?

For example

- WLWW $\rightarrow$ Alice wins
- WLLWLL $\rightarrow$ Alice loses

Observations:

- The game can go on for arbitrarily long
- Special case $p = 1/2$:

Example: Deuce

Alice plays tennis against Bob. The game is at deuce. Suppose

- Alice wins each point with probability p and loses with probability $q = 1 - p$
- Each point is played independently
- The game is won by the player who leads by 2 points

What is the probability that Alice eventually wins the game?

For example

- WLWW $\rightarrow$ Alice wins
- WLLWLL $\rightarrow$ Alice loses

Observations:

- The game can go on for arbitrarily long
- Special case $p = 1/2$: Answer is $1/2$ by symmetry

Example: Deuce

Alice plays tennis against Bob. The game is at deuce. Suppose

- Alice wins each point with probability p and loses with probability $q = 1 - p$
- Each point is played independently
- The game is won by the player who leads by 2 points

What is the probability that Alice eventually wins the game?

For example

- WLWW $\rightarrow$ Alice wins
- WLLWLL $\rightarrow$ Alice loses

Observations:

- The game can go on for arbitrarily long
- Special case $p = 1/2$: Answer is $1/2$ by symmetry
- Instead of listing all the outcomes corresponding to Alice winning, let's do somthing differently...

Let

- $u = P(\text{Alice eventually wins}|\text{game is tied})$
- $v = P(\text{Alice eventually wins}|\text{leading by one point})$
- $w = P(\text{Alice eventually wins}|\text{trailing by one point})$

Let

- $u = P(\text{Alice eventually wins}|\text{game is tied})$
- $v = P(\text{Alice eventually wins}|\text{leading by one point})$
- $w = P(\text{Alice eventually wins}|\text{trailing by one point})$

By law of total probability:

$$\begin{aligned} u =& P(\text{wins game}|\text{draw}) \\ =& P(\text{wins game}|\text{draw}, \text{wins next point}) \times p \\ & + P(\text{wins game}|\text{draw}, \text{loses next point}) \times q \end{aligned}$$

Let

- $u = P(\text{Alice eventually wins}|\text{game is tied})$
- $v = P(\text{Alice eventually wins}|\text{leading by one point})$
- $w = P(\text{Alice eventually wins}|\text{trailing by one point})$

By law of total probability:

$$
\begin{aligned}
u =& P(\text{wins game}|\text{draw}) \\
=& P(\text{wins game}|\text{draw, wins next point}) \times p \\
& + P(\text{wins game}|\text{draw, loses next point}) \times q \\
=& v \times p + w \times q
\end{aligned}
$$

Let

- $u = P(\text{Alice eventually wins}|\text{game is tied})$
- $v = P(\text{Alice eventually wins}|\text{leading by one point})$
- $w = P(\text{Alice eventually wins}|\text{trailing by one point})$

By law of total probability:

$$\begin{aligned}
u =& P(\text{wins game}|\text{draw}) \\
=& P(\text{wins game}|\text{draw, wins next point}) \times p \\
& + P(\text{wins game}|\text{draw, loses next point}) \times q \\
=& v \times p + w \times q
\end{aligned}$$

Similarly,

- $v = 1 \times p + u \times q$
- $w = u \times p + 0 \times q$

Solving the equations

$$\begin{cases} u = vp + wq \\ v = p + uq \\ w = up \end{cases}$$

Solving the equations

$$\begin{cases} u = vp + wq \\ v = p + uq \\ w = up \end{cases}$$

gives

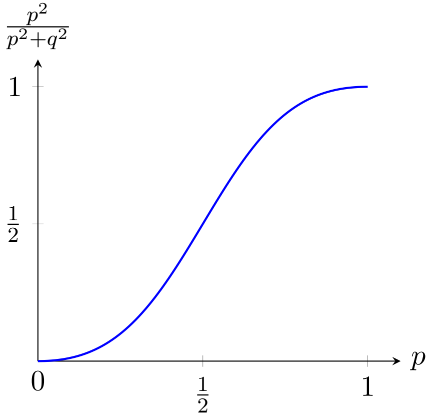
$$u = \frac{p^2}{1 - 2pq} = \frac{p^2}{p^2 + q^2}$$

Solving the equations

$$\begin{cases} u = vp + wq \\ v = p + uq \\ w = up \end{cases}$$

gives

$$u = \frac{p^2}{1-2pq} = \frac{p^2}{p^2+q^2}$$

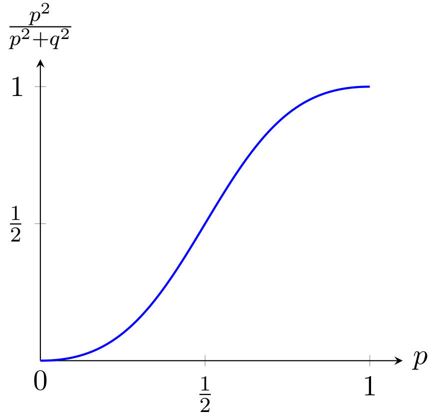


Solving the equations

$$\begin{cases} u = vp + wq \\ v = p + uq \\ w = up \end{cases}$$

gives

$$u = \frac{p^2}{1-2pq} = \frac{p^2}{p^2+q^2}$$



- This method is called **first-step analysis** in B-H Sec 2.7.2
- Will revisit later when discussing random walk

Application: Optimal Stopping

Motel problem

- Suppose you are driving down a highway and pass a strip of n motels

Motel problem

- Suppose you are driving down a highway and pass a strip of n motels
- with distinct rates and arranged randomly

Motel problem

- Suppose you are driving down a highway and pass a strip of n motels
- with distinct rates and arranged randomly
- n is quite large

Motel problem

- Suppose you are driving down a highway and pass a strip of n motels
- with distinct rates and arranged randomly
- n is quite large
- You do not know the arrangement: no :-(

Motel problem

- Suppose you are driving down a highway and pass a strip of n motels
- with distinct rates and arranged randomly
- n is quite large
- You do not know the arrangement: no 📱 :-(
- Once passed a motel, cannot turn back

Motel problem

- Suppose you are driving down a highway and pass a strip of n motels
- with distinct rates and arranged randomly
- n is quite large
- You do not know the arrangement: no  :-(
- Once passed a motel, cannot turn back
- Goal: choose the cheapest motel

A strategy

Phase I: "exploration" Drive past the first k motels and record the prices;

Phase II: "exploitation" Pick the first motel that is cheaper than the cheapest of the first k motels (if none, it's a failure).

A strategy

Phase I: "exploration" Drive past the first k motels and record the prices;

Phase II: "exploitation" Pick the first motel that is cheaper than the cheapest of the first k motels (if none, it's a failure).

Question

- What is $P(\text{success}) = P(\text{find the cheapest})$

A strategy

Phase I: "exploration" Drive past the first k motels and record the prices;

Phase II: "exploitation" Pick the first motel that is cheaper than the cheapest of the first k motels (if none, it's a failure).

Question

- What is $P(\text{success}) = P(\text{find the cheapest})$
- What is the best choice of k?
  - $k = 0$ is too greedy: $P(\text{success}) = \frac{1}{n} \approx 0$
  - $k = n$ is clearly bad: $P(\text{success}) = 0$
  - Maybe some intermediate value k is good

A strategy

Phase I: "exploration" Drive past the first k motels and record the prices;

Phase II: "exploitation" Pick the first motel that is cheaper than the cheapest of the first k motels (if none, it's a failure).

Question

- What is $P(\text{success}) = P(\text{find the cheapest})$
- What is the best choice of k?
  - $k = 0$ is too greedy: $P(\text{success}) = \frac{1}{n} \approx 0$
  - $k = n$ is clearly bad: $P(\text{success}) = 0$
  - Maybe some intermediate value k is good
- How good is this strategy?

Surprisingly, when n is large

- Optimal choice of $k \approx n/e$

Surprisingly, when n is large

- Optimal choice of $k \approx n/e$
- $P(\text{success}) \approx 1/e \approx 36.8\%$

Surprisingly, when n is large

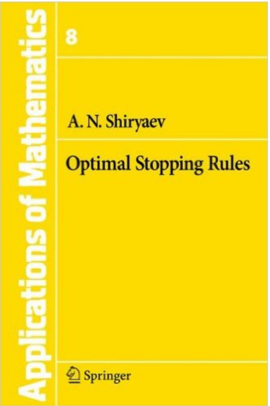
- Optimal choice of $k \approx n/e$
- $P(\text{success}) \approx 1/e \approx 36.8\%$
- This is the best among all strategies (not just two-phase strategies)!!

Surprisingly, when n is large

- Optimal choice of $k \approx n/e$
- $P(\text{success}) \approx 1/e \approx 36.8\%$
- This is the best among all strategies (not just two-phase strategies)!!
- We will use LOTP to resolve the first two questions

Surprisingly, when n is large

- Optimal choice of $k \approx n/e$
- $P(\text{success}) \approx 1/e \approx 36.8\%$
- This is the best among all strategies (not just two-phase strategies)!!
- We will use LOTP to resolve the first two questions
- The third takes a PhD



Mathematically equivalent scenarios

“Best” strategy of the job interviewer:

- Blindly reject the first 36.8% applicants (but keep their CVs)
- Accept the next applicant who beats the best of the first 36.8%

Solution using law of total probability

- Let
$$E_i = \{i\text{th motel is the cheapest}\}, \quad i = 1, \ldots, n$$
Then $P(E_i) = 1/n$.

Solution using law of total probability

- Let
$$E_i = \{i\text{th motel is the cheapest}\}, \quad i = 1, \ldots, n$$
Then $P(E_i) = 1/n$.
- By Law of Total Probability
$$P(\text{success}) = \sum_{i=1}^{n} P(\text{success}|E_i)P(E_i) = \frac{1}{n}\sum_{i=1}^{n} P(\text{success}|E_i)$$

Solution using law of total probability

- Let
$$E_i = \{i\text{th motel is the cheapest}\}, \quad i = 1, \ldots, n$$
Then $P(E_i) = 1/n$.
- By Law of Total Probability
$$P(\text{success}) = \sum_{i=1}^{n} P(\text{success}|E_i)P(E_i) = \frac{1}{n}\sum_{i=1}^{n} P(\text{success}|E_i)$$
- For $i = 1, \ldots, k$,
$$P(\text{success}|E_i) = 0$$

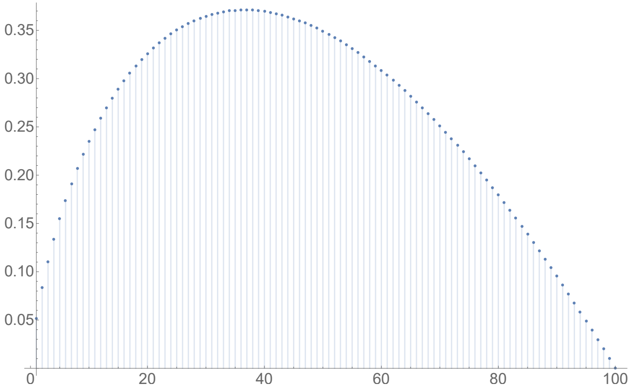
Solution using law of total probability

- Let
$$E_i = \{i\text{th motel is the cheapest}\}, \quad i = 1, \ldots, n$$
Then $P(E_i) = 1/n$.
- By Law of Total Probability
$$P(\text{success}) = \sum_{i=1}^{n} P(\text{success}|E_i)P(E_i) = \frac{1}{n}\sum_{i=1}^{n} P(\text{success}|E_i)$$
- For $i = 1, \ldots, k$,
$$P(\text{success}|E_i) = 0$$
- For $i = k+1, \ldots, n$
$$\begin{aligned} P(\text{success}|E_i) &= P(\text{cheapest of the first } i-1 \text{ is located in the first } k|E_i) \\ &= \frac{k}{i-1} \end{aligned}$$

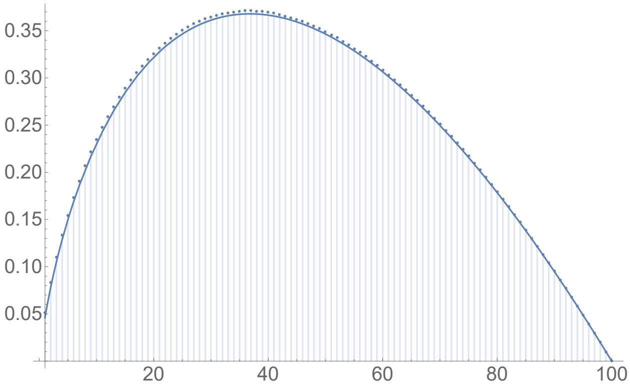
Solution using law of total probability

$$P(\text{success}) = \sum_{i=1}^{k} 0 \times \frac{1}{n} + \sum_{i=k+1}^{n} \frac{k}{i-1} \times \frac{1}{n} = \frac{k}{n} \sum_{i=k}^{n-1} \frac{1}{i}$$

$P(\text{success})$ versus k: $n = 100$



P(success) versus k: $n = 100$



Approximating $\sum$ by $\int$

$$
\begin{aligned}
P(\text{success}) &= \frac{k}{n}\sum_{i=k}^{n-1}\frac{1}{i} \\
&= \frac{k}{n}\left(\frac{1}{n}\sum_{i=k}^{n-1}\frac{n}{i}\right) \\
&\overset{(*)}{\approx} \frac{k}{n}\int_{k/n}^{1}\frac{1}{x}dx = \frac{k}{n}\ln\frac{n}{k}.
\end{aligned}
$$

which attains the maximum $1/e$ at $k = n/e$.

$(*)$: Approximate sum by integtral (see B-H math appendix A.8.4)

Approximating $\sum$ by $\int$

$$
\begin{aligned}
P(\text{success}) &= \frac{k}{n}\sum_{i=k}^{n-1}\frac{1}{i} \\
&= \frac{k}{n}\left(\frac{1}{n}\sum_{i=k}^{n-1}\frac{n}{i}\right) \\
&\overset{(*)}{\approx} \frac{k}{n}\int_{k/n}^{1}\frac{1}{x}dx = \frac{k}{n}\ln\frac{n}{k}.
\end{aligned}
$$

which attains the maximum $1/e$ at $k = n/e$.

$(*)$: Approximate sum by integtral (see B-H math appendix A.8.4)

For more see https://en.wikipedia.org/wiki/Secretary_problem