II T I

Criticate model B.

For model B calculate

\( P(\text{second guesser wins}) \)

Which model A or B do you personally believe is better?

9. PROBABILITIES FOR DICE. DISTRIBUTION OF RANDOM VARIABLE. Two dice, one red and one green, when tossed together give 36 possibilities. Taking the classical model evaluate \( P(R + G = k) \) for each of \( k = 2 \) through 14. These probabilities comprise the "probability distribution of \( R+G \)."

<table>
<thead>
<tr>
<th>red</th>
<th>green</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\[ P(R + G = 5) = \frac{4}{36} = \frac{1}{9} \]

\( P(R + G = k) \)

Later on we will formally introduce the notion of "random variable" which is simply a numerical outcome of a probability experiment. Each of \( R, G, R+G \) would be a random variable on the experiment of tossing two dice. From the distribution of \( R+G \) calculate

\( P(R + G < 4) \)

\( P(|R+G - 7| > 1) \)

10. TWO DICE VS ONE DIE. Calculate

\( P(R = 5) \) from the classical model for the toss of the red die alone

\( P(R = 5) \) from the classical model for the toss of two dice.

See that these two calculations agree.
10. **NON-TRANSITIVE DICE, NO NATIONAL CHAMP.** The integer numbers 1 through 24 are assigned to the faces of four cubes to produce special dice as follows (B. Efron, with different numbers)

<table>
<thead>
<tr>
<th>die A</th>
<th>10</th>
<th>1</th>
<th>12</th>
<th>13</th>
<th>14</th>
<th>15</th>
</tr>
</thead>
<tbody>
<tr>
<td>die B</td>
<td>1</td>
<td>2</td>
<td>16</td>
<td>17</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>die C</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>20</td>
<td>21</td>
<td>22</td>
</tr>
<tr>
<td>die D</td>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
<td>23</td>
<td>24</td>
</tr>
</tbody>
</table>

Using the methods of (7) calculate each of

P(die D > die C)

P(die C > die B)

P(die B > die A)

P(die A > die D)

Here is a game. A person chooses one of these dice. A second person chooses any die they wish from the three remaining. These two dice are then tossed and the one showing the higher number wins.

Which die would you choose if I first chose die B?

Should you play first or second in the game?

Is there a "best" die?

Is there necessarily a best football team, even if very many games are played to settle it?

11. **JACK AND JILL, CONFRONTING OUR SOMETIME UNEASE WITH PROBABILITY CALCULATIONS.** A box \[ \{a, b, c, 1, 5, 5\} \] holds three bills of which one is a five and two are ones. Jack will select a bill at random (classical model). So \( P(Jack = 5) = \frac{1}{3} \). Jill will then select a bill with equal probability from the two bills remaining. What is \( P(Jill = 5) \)?

A classical model for the selection of two bills, without replacement, is needed to answer this question. We are drawing bills, not dollar amounts, so it is useful to consider the sample space to be all 6 possible draws of two, without replacement, from \( \{a, b, c\} \) where "c" denotes the $5 bill.

<table>
<thead>
<tr>
<th></th>
<th>jack</th>
<th>jill</th>
<th>jack = 5</th>
<th>jill = 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>a</td>
<td>b</td>
<td>no</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>c</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>a</td>
<td>no</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>b</td>
<td>c</td>
<td>no</td>
<td>yes</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>a</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
<tr>
<td>c</td>
<td>b</td>
<td>yes</td>
<td>no</td>
<td></td>
</tr>
</tbody>
</table>

By enumerating these six possibilities and checking off favorable cases calculate
P(Jill | 5)

P(Jack | 5).

This example illustrates the principle that order of the deal does not matter even when cards are not returned to the deck. Although you may have had trouble with guessing P(Jill | 5) you surely act in accordance with the order of the deal principle. After all, people are not seen to quarrel over who gets to sit next to the dealer at cards.

1.2. CONDITIONAL PROBABILITY PRINCIPLE FOR BLACKJACK, NOT ALL CARD GAMES. In blackjack, we get to see cards dealt face up to the table. The principle says that we may then revise our probabilities by applying the classical model to an imagined deck from which are removed all of the cards that have been seen. This does not apply to cards seen through other means than a random deal, such as may occur in other card games or by peeking. Understanding why helps us understand conditional probability.

For example, as shown below, if I hold the ace of diamonds and the seven of hearts and nothing else is known then the probability that the ace of spades is held by a given player holding two cards is \( \frac{1 \times 49}{50 \times 49} = \frac{1}{25} \).

**BRUTE FORCE, WITHOUT THE PRINCIPLE.** To appreciate this principle we have to imagine life without it. In order to calculate the above probability by brute force, without using the imagined deck shortcut, we would need to begin with a sample space for all possible “table layouts.” A table layout is a possible layout of all of the cards that have been dealt to all of the players. We would then remove from this sample space any table layout inconsistent with the cards we have seen. So if I have only seen the ace of diamonds and the seven of hearts I would strike off all layouts that were inconsistent with this. If I learned about other cards because I saw them dealt to the table I would strike off all layouts inconsistent with that information. For this reduced sample space of table layouts the classical probability that a given player holds the ace of spades among their two cards is

\[
\text{P(the given player holding two cards has the ace of spades) =}
\frac{\text{number of table layouts consistent with cards we have seen in which the given player holds the ace of spades}}{\text{total number of possible table layouts consistent with cards we have seen}}.
\]

**THIS WOULD BE HARD! THE PRINCIPLE FOR BLACKJACK SAYS WE DON’T HAVE TO DO ALL OF THAT.** We simply imagine that the player of interest is being dealt two cards from a deck from which have been removed all cards we have seen.

**APPLYING THE PRINCIPLE.** We can make this calculation by imagining her two cards laid side by side. We will choose the sample space of all possible left-right pairs of cards. So, for example, the contingencies (ace of spades, 6 of clubs) and (6 of clubs, ace of spades) are distinct points of this sample space.

Keep in mind that the ace of diamonds and seven of hearts have been removed from the imagined deck of 52. Her left card may thus be any of 50 and, no matter what it is, her right card may be any of 49. So there are 50\times49 possible ordered pairs for her two cards. This the number of points in our sample space.

We must now count the ways she can have an ordered pair of which one is the ace of spades. If she has the ace of spades in her left place there are 49 possible cards for the right place. Similarly, if she has the ace of spades in the right place there are 49 possible cards for her left place. So there are only 2\times49 left-right pairs that hold the ace of spades.

To re-cap, if we know that the ace of diamonds and seven of hearts are not in her hand of two cards, we quote the probability \( \frac{2 \times 49}{50 \times 49} = \frac{1}{25} \) that she has the ace of spades.

**EXERCISE.** Once you have understood this way of thinking do the following: A deck has cards 1, 2, 3, 4 from which I am dealt the two cards 1 and 4. You were dealt a single card. Calculate

P(your card is 2 | it is known that my cards are 1 and 4)

Later on, we will use the notation P(your card is 2 | my cards are 1 and 4) where the vertical bar “|” is read “given that.” This is called a “conditional” probability.

**BUT NOT FOR ALL CARD GAMES.** We do have to sharpen the statement of the Principle given above. If I had peeked at the other player’s hand and seen (say) a ten of diamonds then the chance she has the ace of spades for her other card is \( \frac{48}{49} \) if we
assume I pecked at one of her two cards at random. But if I pecked and saw her upper card in a fan arrangement of the two I might be justified in thinking it is her high card since it is common for players to arrange cards in that order of size. In such a case I would quote P(she has the ace of spades) = 0. In some other card games we may see cards because a player does not want them. This introduces a lot of subtleties as to what kind of information we are getting and how that should revise our probabilities. Poker is for thinkers! The above Principle does not apply to poker except when you see cards at random.

13. CONDITIONAL PROBABILITY. Refer to (1.1) and (1.2) above. Calculate

\[ P(\text{Jill} \mid \text{Jack}) \text{, i.e. } P(\text{Jill} \mid \text{Jack}) \]

\[ P(\text{Jack} \mid \text{Jill}) \]

14. LET’S MAKE A DEAL, WHAT IS A GOOD MODEL?

The test of any model is whether we can discover anything about it that doesn’t seem faithful to the circumstances or that goes beyond the knowledge we have.

Mr. Barker has placed a prize behind one of three doors. There is nothing of value behind the other two doors. You are invited to “choose a door.” If you have chosen your door at random (equal chance for each door) then P(you have chosen the winning door) = 1/3. A classical model for the random selection of a door is

model A

\[ \begin{array}{ccc}
1 & 2 & 3 \\
\end{array} \]

Mr. Barker now opens one of the other two doors showing that there is no prize behind it. He then offers you the option of switching from your choice to the unopened door that you did not choose at first. Should you do it? Notice that no matter where the prize happens to be you always win by switching doors unless you have originally chosen the correct door. Intuit

P(you win by switching)

Model A may not be detailed enough to give us confidence in the answer. In particular, what is the subset of the classical model that is identified with winning by switching? Is there one?

15. LET’S MAKE A DEAL CONT. Consider another proposal, this one for a very explicit sample space for the problem:

model B

\[ \begin{array}{cccc}
112 & 113 & 123 & 132 \\
221 & 223 & 213 & 231 \\
331 & 332 & 312 & 321 \\
\end{array} \]

Notation 132 here means that the prize is behind door 1, you choose door 3, and Mr. Barker reveals that door 2 has no prize behind it. In that case switching from 1 to 3 will get you the prize. The model seems impressively detailed.

In model B circle each case in which switching will win and calculate

P(you win by switching)

This is not the answer we got for model A, so something may be amiss. Observe that Model B makes P(prize is behind 1 and you choose 1) = 2/12, twice the size of P(prize is behind 1 and you choose 2) = 1/12. This is not consistent with our choice of a door at random.

Yet another model balances that inequity

model C

\[ \begin{array}{cccc}
112 & 113 & 123 & 132 & 132 \\
\end{array} \]
Sure enough, for this new sample space \( P(\text{prize is behind 1 and you choose 1}) = \frac{2}{18} = P(\text{prize is behind 1 and you choose 2}) \). For model C calculate

\[
P(\text{you win by switching})
\]

This agrees with our answer from model A so perhaps this model is correct.

There are some things about model C we do not like however! It says \( P(\text{prize is behind door 2}) = \frac{6}{18} = \frac{1}{3} \). How could we know such a thing? What if Mr. Barker favors placing the prize behind door one? We don’t know anything about how he is making his choices. Model C seems reasonable and gives an agreeable answer but it goes far beyond what we know for sure.

16. LET’S MAKE A DEAL, CONT.

Return to the original classical model for our choice of a door at random

model C

1 \hspace{1cm} 2 \hspace{1cm} 3

For this model we cannot calculate \( P(\text{you win by switching}) \) because we don’t know where the prize has been placed. But we can calculate it for each of the three cases: prize behind door 1, door 2, door 3. You will find they are all the same.

\[
P(\text{you win by switching if prize behind door 1}) \]

\[
P(\text{you win by switching if prize behind door 2})
\]

\[
P(\text{you win by switching if prize behind door 3})
\]

Your answer happens to be the same as you intuit in (15) but without assuming anything about how Mr. Barker has chosen a door behind which to put the prize.

By randomly choosing a door we have injected probability into the problem. Then we are in a position to quote odds. This is the advantage behind heavy uses of randomization such as choosing a random sample of customers.

Does there remain anything you don’t like about model A?

17. LET’S MAKE A DEAL CONT. From (14) through (16) you are seeing how the choice of a probability model can affect the answer and that one way you challenge a model is to look to some of its features. You are also seeing that there may be sensible ways to choose a model based upon the circumstances and that “randomization” can be a powerful tool for justifying probability calculations and getting practical answers. How would you play a round of Let’s Make A Deal?

18. SALESPERSONS’ PROBLEM. Customers enter a retail outlet in random order one by one. Sales people wait to take a customer. Suppose that you are able to choose the point at which you step in to take a customer. You will do so when a customer looks good to you, especially since there are occasionally customers who earn you a lot of money. We don’t know how much each customer will earn us in commission but we are able to judge how they appear. As an example, suppose there are four arriving customers (which makes for \( 24 = 4! \times 2 \times 1 \) possible arrangements as to how the customers will rank in your eyes). A sample space is

\[
\begin{align*}
1234 & \quad 1243 & \quad 1324 & \quad 1342 & \quad 1423 & \quad 1432 \\
2134 & \quad 2143 & \quad 2314 & \quad 2341 & \quad 2413 & \quad 2431
\end{align*}
\]
9124 9142 9214 9241 9412 9421
4123 4132 4213 4231 4312 4321
where 3142 means that your third best prospect will arrive first followed by the one you rank first, etc. Your problem is that before you get to see the next customer another salesperson may snatch the one who just entered, so you have to act fast based upon what you have seen so far.

We'll compare three strategies that you might adopt. The first one is to simply take whoever walks in first. For this strategy P(you select your rank 1 customer) = \frac{6}{24} = \frac{1}{4}.

Another strategy would be to look at the first customer to enter but not take them. You could then take the next customer to subsequently enter who beats the first one, or just take customer 4 if that fails to happen. By counting cases calculate

P(you select rank 1 customer)

for this "pass up the first" strategy.

Suppose that you instead size up the first two entering customers and take the first subsequent customer to best them (or customer 4 if that fails to happen)? Evaluate

P(you select rank 1 customer)

for this "pass up the first two" strategy. Based on this analysis which do you prefer to do?

19. SALESPEOPLE'S PROBLEM CONT. If you take the customer who enters first then

<table>
<thead>
<tr>
<th>rank of customer selected</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>probability</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
</tr>
</tbody>
</table>

is the "distribution of the rank of the selected customer."

Find the distribution of rank of the selected customer for the strategy of "take the first subsequent customer better than the first"

<table>
<thead>
<tr>
<th>rank of customer selected</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>probability</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
</tr>
</tbody>
</table>

Find the distribution also for the strategy "take the first subsequent customer better than the first and second."

<table>
<thead>
<tr>
<th>rank of customer selected</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
</tr>
</thead>
<tbody>
<tr>
<td>probability</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
<td>\frac{1}{4}</td>
</tr>
</tbody>
</table>

At yet another, but large, dinner gathering of professionals I found myself passing along pieces of pie as they were cut by our hostess Ms. T. W. Anderson. Somebody caught on and we had a few laughs over this application of the salespersons' problem (then called the secretary problem). Along came a monster piece of pie which of course was the one I chose not to pass. Ms. Anderson had surely heard of the problem and was having a little fun too.

20. NUMBER e. The irrational number e = 2.7182818284590... comes into play when we compound many times. It has earned an illustrious role in many fields, entering naturally when many tiny numbers are multiplied. It is rooted in the fact, not proven here, that \( \left(1 + \frac{x}{n}\right)^n \) quickly approaches \( e^x \) as \( n \to \infty \), for every real number \( x \). You may define \( e \) as the limit for \( x = 1 \).

You are offered CD-I which returns 5% annual, compounded monthly. Another CD-II earns 4.8% annual compounded daily. A third earns 4.77% annual, compounded continuously. What does $1 invested in each product over six months? Which has the best return? Solve this after you read what follows.

If a bank offers CD-I at "5% annual, compounded monthly" what they mean is that after \( n \) months each dollar has grown to
\[(1 + \frac{0.05}{12})^6 = 1.004167 \approx 1.004167\]. For example, after \(n = 6\) months one dollar will grow to exactly 
\[1.004167 = \frac{1.000000}{1.000000} = 1.025266\] (decimal expansion stopped but not rounded).

Using the e-approximation we come pretty close, obtaining 
\[(1 + \frac{0.05}{12})^6 = (1 + \frac{0.05}{12})^{12} \approx e^{0.05} \approx 1.05127.

The e-approximation simply stated: \((1 + \text{tiny})^\text{large}\) is approximately \(e^\text{tiny} \cdot \text{large}\). This is a useful approximation to keep in mind\(^1\).

Another bank offers \(CD-I\) at 4.8\% annual compounded daily. For the \(155 = 31 + 29 + 31 + 30 + 31\) days of January through June (in a leap year) a dollar grows to \(1 + \frac{0.048}{365} = 1.0205913\), the exponential approximation of this being \(e^{0.048/365} = 1.0205927\).

\(CD-III\) offers "continuous compounding" by which they mean the limit of compounding by the second, by the nanosecond, and so forth. This just means they will pay the e-approximation! So after 155 days of continuous compounding one dollar will grow to \(e^{0.048/365} = 1.0205927127498\). As a check, compounding by the second, at 4.77\% annual rate of return, over 155 days, one dollar will grow to 
\[(1 + \frac{0.0477}{365})^{155} = 1.0204627127342\] requiring minutes to compute. We've assumed the bank continues to work with 365 days even though it is a leap year.

From these calculations we see that (1) the e-approximation is surprisingly accurate in these applications and (2) \(CD-I\) offers the highest return over the 6 months.

\(^1\) From calculus, \((1 + x)^r = e^{x} = x + \left(x - \frac{x^2}{2}\right) + O\left(\frac{1}{y}\right)^3 + O(x)^3\) as \(x \to 0, y \to \infty\) (a fine point).

21. SALESPEOPLE'S PROBLEM CONT. The salespeople's problem has many applications to situations where one must act promptly or lose the opportunity: selecting a mate, a house, an item on eBay, a trophy animal or fish, a job offer, a proposal, an investment, etc.

An interviewer plans to meet with each of 17 candidates for a position until that position is filled "on the spot" by hiring the person currently being interviewed. This scheme has been adopted because virtually all candidates who become available find a job immediately and the company thinks it can hold out through 17 interviews given the rate at which candidates become available. It is felt that candidates arrive in random order as regards their quality. Based on the ideas of (20) a strategy has been put forth to: (a) interview, but not hire, \(\frac{17}{e} = 6\) candidates then (b) hire the interviewee who appears better than those 6 (hiring the 17th if that fails to happen). It is claimed (not proven here) that we hire the very best candidate of all 17 with probability around \(\frac{1}{e} \approx 0.367879\) using this strategy.

For your exercise, try this by making up 17 cards numbered 1 through 17. Shuffle them. Examine the first 6, noting the lowest number. Continue looking at the remaining cards one by one until a still lower number is seen (the hire) or you have to settle for the 17th card selected. Repeat this activity many times. You should hire number 1 in around 38\% of cases. Does it seem to work?

For a million interviews we would examine the first \(\frac{1000000}{e} = 367,879\), noting the best of them, then continue through the rest of the million interviews, hiring the first one who is better than all previous (or the last if none appears). This strategy has around 0.368 probability of hiring the best of the one million, at least the one we judge to be the best at the time of hiring. Instead of interviews think of a computer scanning segments of DNA looking for a "best match."