Interesting Probability Questions[2]

Some Birthday and Lottery Problems[3]

1. Assume that there are *n* people in the room. Ignoring leap years, what is the mathematical formula for the probability that <u>no</u> one else in the room shares your birthday?
2. Assume that there are 253 people in the room. Ignoring leap years, what is the probability that no one else in the room shares your birthday? What is the probability that <u>at</u> <u>least</u> one other person in the room shares your birthday?
3. Assume that there are *n* people in the room. Ignoring leap years, what is the mathematical formula for the probability that <u>no</u> two people share the same birthday?
4. Assume that there are *n* people in the room. Ignoring leap years, what value of *n* makes the probability that <u>at</u> <u>least</u> two people in the room share birthdays equals .5 (most closely)?
5. Why is the number of people in question number 2 larger than the number of people in question 4?
6. In the North Carolina Lottery *Cash-5 Game*, the same five numbers [4, 21, 23, 24, 39] were randomly selected on July 9, 2007, and July 11, 2007. Given that there are 39 numbers that can be selected without replacement, what is probability of <u>at</u> <u>least</u> one matching set of 5 numbers in a given set of 3 drawings?

The Medical Diagnosis Problem[4]

7. Assume that a test to detect a disease whose prevalence is (1/1000) has a false positive rate of 5% and a true positive rate of 100%. What is the probability that a person found to have a positive result actually has the disease assuming that you know nothing about the person's symptoms?

"Bertrand's Box" Problem[5]

8. A box has three drawers; one contains two gold coins, one contains two silver coins, and one contains one gold coin and one silver coin. Assume that one drawer is selected randomly and that a randomly selected coin from that drawer turns out to be gold. What is the probability that the chosen drawer is one that contains two gold coins?

"Let's Make A Deal" Problem[6]

9. Suppose that you're on a game show and you're given the choice of three doors. Behind one door is a car, and goats are behind the other two doors. You pick a door and the host, who knows what behind each door, opens another door to show a goat. He says, "Do you want to switch doors?" Should you switch?

---

[2] All of these questions are common ones, but I took the suggestion to use them from: Richard G. Seymann. November 1991. Comment. *The American Statistician*. **45**(4):287-8. **The answers are linked to my web page.**

[3] Questions 1-5 are due to: William Feller. 1957. *An Introduction to Probability Theory and its Applications (Volume 1).* New York: John Wiley. Question 6 is based on: Leonard A. Stefanski. May 2008. The North Carolina Lottery Coincidence. *The American Statistician*. **62**(2):130-134.

[4] This question was presented to 60 students and staff at the Harvard Medical School. Over half of the participants responded 95% and only 11 gave the correct response. Hmm… See: D. Kahneman, P. Slovik, and A. Tversky. 1982. *Judgement Under Uncertainty: Heuristics and Biases*. Cambridge, UK: Cambridge University Press.

[5] See: R. Weatherford. 1982. *Philosophical Foundations of Probability Theory*. London: Routledge and Kegan Paul.

[6] The firestorm includes: vos Savant, Marilyn. September 9, 1990. "Ask Marilyn". *Parade*. 15; December 2, 1990. "Ask Marilyn". *Parade*. 25; February 17, 1991. "Ask Marilyn". *Parade*. 15.
Morgan, J.P., *et al*. November 1991. "Let's Make a Deal: The Player's Dilemma". *The American Statistician*. **45**(4):284-7.
Seymann, Richard G. November 1991. "Comment". *The American Statistician*. **45**(4):287-8.
Morgan, J.P., *et al*. November 1991. "Rejoinder". *The American Statistician*. **45**(4):289.
vos Savant, Marilyn. November 1991. "Comment". *The American Statistician*. **45**(4):347.
Morgan, J.P., *et al*. November 1991. "Rejoinder to vos Savant". *The American Statistician*. **45**(4):347-348.

Some Birthday and Lottery Problems

1. Assume that there are *n* people in the room. Ignoring leap years, what is the mathematical formula for the probability that no one else in the room shares your birthday?

   The mathematical formula is: $p = \left(\frac{364}{365}\right)^n$

2. Assume that there are 253 people in the room. Ignoring leap years, what is the probability that no one else in the room shares your birthday? What is the probability that at least one other person in the room shares your birthday?

   Since: $p = \left(\frac{364}{365}\right)^{253} = .499523 \approx .5$

   Therefore: $1 - \left(\frac{364}{365}\right)^{253} = 1 - .499523 \approx .5$

3. Assume that there are *n* people in the room. Ignoring leap years, what is the mathematical formula for the probability that no two people share the same birthday?

   The formula is: $p = \left(\frac{364}{365}\right)\left(\frac{363}{365}\right)\left(\frac{362}{365}\right)\ldots\left(\frac{365-(n-1)}{365}\right) = \frac{(365-1)(365-2)\ldots(365-(n-1))}{365^{n-1}}$

4. Assume that there are *n* people in the room. Ignoring leap years, what value of *n* makes the probability that at least two people in the room share birthdays equals .5 (most closely)?

   Let n=23: $1 - p = 1 - \left(\frac{364}{365}\right)\left(\frac{363}{365}\right)\left(\frac{362}{365}\right)\ldots\left(\frac{365-(23-1)}{365}\right) = 1 - .492703 = .507297 \approx .5$

5. Why is the number of people in question number 2 larger than the number of people in question 4?

   Question 2 asks about a specific date (your birthday) while question 4 asks about any of the possible 365 days.

6. In the North Carolina Lottery *Cash-5 Game*, the same five numbers [4, 21, 23, 24, 39] were randomly selected on July 9, 2007, and July 11, 2007. Given that there are 39 numbers that can be selected without replacement, what is probability of at least one matching set of 5 numbers in a given set of 3 drawings?

$$p = 1 - \frac{(k-1)(k-2)\ldots(k-(n-1))}{k^{n-1}} = 1 - \frac{(k-1)(k-2)}{k^{3-1}} = 1 - 0.999994789 = 5.211x10^{-6} \approx \frac{1}{191{,}919}$$

$$k = {}_{39}C_5 = \frac{39!}{5!34!} = 575{,}757$$

Note that this is the same basic concept as question 4.

7. Assume that a test to detect a disease whose prevalence is (1/1000) has a false positive rate of 5% and a true positive rate of 100%. What is the probability that a person found to have a positive result actually has the disease assuming that you know nothing about the person's symptoms?

D = Has the disease
T = Test result is positive
Given: $P(D) = .001$ so $P(\text{Not } D) = .999$
$P(T \mid \text{Not } D) = .05$
$P(T \mid D) = 1.00$
$P(D \mid T) = P(T \cap D) / P(T)$
$P(T \cap D) = P(D)P(T \mid D)$
$= .001$
$P(T) = P(T \cap D) + P(T \cap \text{Not } D)$
$= P(T \cap D) + P(\text{Not } D)P(T \mid \text{Not } D)$
$= .001 + (.999)(.05)$
$= .05095$
$P(D \mid T) = .001/.05095 = .019627$

8. A box has three drawers; one contains two gold coins, one contains two silver coins, and one contains one gold coin and one silver coin. Assume that one drawer is selected randomly and that a randomly selected coin from that drawer turns out to be gold. What is the probability that the chosen drawer is one that contains two gold coins?

Initial set-up

1. Assuming that the boxes are selected randomly:
   A = Box A (2 gold coins) P(A) = P(Pick Box A) = 1/3
   B = Box B (1 gold & 1 silver coin) P(B) = P(Pick Box B) = 1/3
   C = Box C (2 silver coins) P(C) = P(Pick Box C) = 1/3

2. Since there are six coins, 3 of each type, and any one coin can be selected randomly:
   P(G) = P(Pick a Gold Coin) = P(Pick a Silver Coin) = P(Not G) = ½

3. Given (1) and (2), the marginal probabilities on the outside of the joint probability table must be:

	Pick a Coin		
Pick Box	**Gold**	**Silver**	**P(Pick Box)**
Box A			P(A)=1/3
Box B			P(B)=1/3
Box C			P(C)=1/3
P(Pick Coin)	P(G)=1/2	P(Not G)=1/2	1

4. Further, given the description of each box in (1) the joint probability table must be the following:

	Pick a Coin		
Pick Box	**Gold**	**Silver**	**P(Pick Box)**
Box A	1/3	0	P(A)=1/3
Box B	1/6	1/6	P(B)=1/3
Box C	0	1/3	P(C)=1/3
P(Pick Coin)	P(G)=1/2	P(Not G)=1/2	1

$P(A \mid G) = P(A \cap G)/P(G) = (1/3)/(1/2) = 2/3$

9. Suppose that you're on a game show and you're given the choice of three doors. Behind one door is a car and behind the other two doors are goats. You pick a door and the host, who knows what behind each door, opens another door to show a goat. He says, "Do you want to switch doors?" Should you switch?

<u>Initial set-up assuming that you've pick Door #1 and before the host shows you what's behind another door</u>

1. Assuming that the car can be behind any one of the three doors and that the host is equally likely to show you what's behind either Door #2 or Door #3, the marginal probabilities on the outside of the joint probability table must be the following:

	Host Shows You Door		
Car Behind Door	**Door #2**	**Door #3**	**P(Car Behind Door)**
Door #1			P(Car D1)=1/3
Door #2			P(Car D2)=1/3
Door #3			P(Car D3)=1/3
P(Host Shows Door)	P(Shows D2)=1/2	P(Shows D2)=1/2	1

2. The car either is as likely to be behind the door that you picked as it is to not be behind the door that you picked, so one set of joint probabilities must be:

	Host Shows You Door		
Car Behind Door	**Door #2**	**Door #3**	**P(Car Behind Door)**
Door #1	1/6	1/6	P(Car D1)=1/3
Door #2			P(Car D2)=1/3
Door #3			P(Car D3)=1/3
P(Host Shows Door)	P(Shows D2)=1/2	P(Shows D2)=1/2	1

3. If the host shows you what's behind door #2, then the car must not be behind it. If the host shows you what's behind door #3, then the car must not be behind it. So, the remaining set of joint probabilities must be:

	Host Shows You Door		
Car Behind Door	**Door #2**	**Door #3**	**P(Car Behind Door)**
Door #1	1/6	1/6	P(Car D1)=1/3
Door #2	0	1/3	P(Car D2)=1/3
Door #3	1/3	0	P(Car D3)=1/3
P(Host Shows Door)	P(Shows D2)=1/2	P(Shows D2)=1/2	1

Now, the host shows you what's behind a door (assume that it's door #2) and asks, "Do you want to switch doors (from door #1 to door #3)?"

P(Car Behind Door #1 | Host Shows Door #2) = P(Car D1 | Shows D2)
= P(Car D1 ∩ Shows D2)/P(Shows D2)
= (1/6)/(1/2)
= 1/3

P(Car Behind Door #3 | Host Shows Door #2) = P(Car D3 | Shows D2)
= P(Car D3 ∩ Shows D2)/P(Shows D2)
= (1/3)/(1/2)
= 2/3

Since P(Car Behind Door #1 | Host Shows Door #2) < P(Car Behind Door #3 | Host Shows Door #2), you should switch from door #1 to door #3.