

Problem 1

You flip 12 coins, what are the odds that 7 of them are heads?

$$\binom{12}{7} \cdot .5^7 \cdot (1 - .5)^{12-7} = .193$$

Problem 2

What are the odds of rolling 7 with 2 dice?

$$6 \cdot \frac{1}{6 \cdot 6} = \frac{1}{6}$$

Problem 3

What is the probability of drawing a pair in a 5 card hand?

Note: for this question 3 of a kind, 2 pair, and 4 of a kind will all count as pairs

Find the compliment of all the cards having different values:

$$1 - \frac{52}{52} * \frac{48}{51} * \frac{44}{50} * \frac{40}{49} * \frac{36}{48} = .49$$

Problem 4

You flip a coin until you get heads, how many times would you expect to flip the coin?

2 times

(a) What is the probability of getting heads on the n^{th} flip?

Assuming this is the first heads:

$$\frac{1}{2^n}$$

(b) What is the expected value of this distribution over all n ?

$$\frac{1}{2}$$

Note: We haven't learned enough to actually solve this problem, so you don't need to find a closed form. However, if you can find a closed form, then feel free.

Problem 5

Assuming that an M&Ms bag has exactly 13 of each color. What are the odds that you pull out 3 red M&Ms from the bag.

$$13 \cdot 6 = 78 \text{ M\&Ms total}$$

Without replacement:

$$\frac{13}{78} \cdot \frac{12}{77} \cdot \frac{11}{76} = \frac{1}{266}$$

Problem 6

When buying a lottery tick you notice that the payouts are a little different. Each ticket costs \$1, and you pick 6 out of the 100 numbers. You get \$1 for matching 1 number, \$10 for matching 2 numbers, and so on up to \$100,000 for matching all 6 numbers. What is the expected payout of this lottery?

There are $\binom{100}{6}$ ways to choose the 6 numbers.

For one matching number there are $\binom{6}{1}$ ways to choose the single matching number and $\binom{94}{5}$ ways to choose the wrong numbers. This pattern continues for the rest of the winnings classes

one number	$\frac{\binom{6}{1} \cdot \binom{94}{5}}{\binom{100}{6}} = .276$
two numbers	$\frac{\binom{6}{2} \cdot \binom{94}{4}}{\binom{100}{6}} = .0383$
three numbers	$\frac{\binom{6}{3} \cdot \binom{94}{3}}{\binom{100}{6}} = .00224$
four numbers	$\frac{\binom{6}{4} \cdot \binom{94}{2}}{\binom{100}{6}} = .000055$
five numbers	$\frac{\binom{6}{5} \cdot \binom{94}{1}}{\binom{100}{6}} = 4.73 \cdot 10^{-7}$
six numbers	$\frac{\binom{6}{6} \cdot \binom{94}{0}}{\binom{100}{6}} = 8.38 \cdot 10^{-10}$

To find the expected value of this lottery each of these probabilities must be multiplied by the payout and summed. Represented as:

$$\sum_{i=0}^5 \frac{\binom{6}{1+i} \cdot \binom{94}{5-i} \cdot 10^i}{\binom{100}{6}} = .9447 \text{ dollars}$$

Problem 7

There is a new drug out that is advertised as curing boringness. Obviously this is a huge breakthrough in medical science. And so far the results seem promising. If you are boring, there is an 80% chance that the

drug will make you more interesting. However, there's a catch. If you're not boring, then there's an 80% chance that the drug will make you less interesting. It's well known that most people are boring. We'll estimate this by saying that 90% of people are boring. What is the probability that this drug will backfire, and make you less interesting? remember: you don't know if you're boring or not.

For the drug to backfire you have to start with being interesting. This is a .1 probability. Then there is a .8 probability the drug backfires.

$$.1 \cdot .8 = .08$$

Problem 8

Oh NO!

We gave the drug from the last problem to little Timmy, and he immediately asked if he could read Moby Dick. Clearly the drug backfired.

I should really stop picking on classic novels.

What is the probability that Timmy boring to begin with?

Let's define a state for the effect of the drug to create some conditional probabilities. This state is success and failure. If you are boring, the drug succeeds if it makes you more interesting, but fails if it doesn't. If you are interesting, it fails if it makes you less interesting but succeeds if it doesn't.

Here are the given probabilities. Probability of:

being boring, $P(B) = .9$

not being boring, $P(B^c) = .1$

drug succeeding if you are boring, $P(S | B) = .8$

drug failing if you are boring, $P(S^c | B) = .2$

drug succeeding if you are not boring, $P(S | B^c) = .2$

drug failing if you are not boring, $P(S^c | B^c) = .8$

Okay, the question is asking for the probability that Timmy is boring given that the drug failed (because if it worked he would be playing video games). WANT: $P(B | S^c)$

$$\text{By Bayes' theorem: } P(B | S^c) = \frac{P(S^c | B) \cdot P(B)}{P(S^c)}$$

$$\text{And by the Law of Total Probability } P(S^c) = P(S^c | B) \cdot P(B) + P(S^c | B^c) \cdot P(B^c)$$

$$\text{Substitute } P(B | S^c) = \frac{P(S^c | B) \cdot P(B)}{P(S^c | B) \cdot P(B) + P(S^c | B^c) \cdot P(B^c)}$$

$$= \frac{(.2)(.9)}{(.2)(.9) + (.8)(.1)}$$

$$P(B | S^c) = .69$$

Problem 9

Prove that the expected value of a random variable is linear. That is, prove $E[aX + b] = aE[X] + b$.

$$E(X) = \sum_{i=1}^n x_i P(X_i) \quad \text{definition of expected value}$$

$$\sum_{i=1}^n P(X_i) = 1 \quad \text{definition of random variable}$$

$$E(aX + b) = \sum_{i=1}^n (ax_i + b)P(X_i) \quad \text{definition of expected value}$$

$$= a \sum_{i=1}^n x_i P(X_i) + \sum_{i=1}^n P(X_i)b \quad \text{algebra}$$

$$= aE(X) + b \quad \text{using definition of expected value and random variable}$$

Problem 10

Veritasium made a video “Is Most Published Research Wrong”.

<https://www.youtube.com/watch?v=42QuXLuch3Q>

In this video he goes over a topic that we don't really talk about a lot in science. It is possible to do everything correctly, and still get wrong result published surprisingly often. He then goes on to other statistical problems that can arise in research.

For this problem I want to take a look at the importance of replication studies. These are studies where a scientist doesn't do anything other than verify other scientists' work. These can reduce the problems raised in this video significantly. In the original video, he shows how about 1/3 of published results could be incorrect using Bayes' theorem. Use Bayes' theorem to find the odds of a study being incorrect after two experiments verifying it.

First let's list the probabilities Veritasium uses to get to that $\frac{1}{3}$ number to make it formal.

Probability of: a conjecture being correct, $P(C) = .1$

a conjecture being incorrect, $P(C^c) = .9$

a correct conjecture resulting in a true positive (statistical power) $P(+ | C) = .8$

a correct conjecture resulting in a false negative $P(+^c | C) = .2$

an incorrect conjecture resulting in a false positive (p-value) $P(+ | C^c) = .05$

an incorrect conjecture resulting in a true negative $P(+^c | C^c) = .95$

Now let's check Veritasium's math on the probability of a conjecture being correct given a positive hypothesis

test.

$$\begin{aligned}
 \text{Want probability of C given +} \quad & P(C \mid +) = \frac{P(+ \mid C) \cdot P(C)}{P(+)} \\
 \text{By the law of total probability} \quad & P(+)=P(+ \mid C) \cdot P(C)+P(+ \mid C^c) \cdot P(C^c) \\
 \text{so...} \quad & P(C \mid +)=\frac{.8 \cdot .1}{.8 \cdot .1+.05 \cdot .9} \\
 & =.64
 \end{aligned}$$

Good it looks like we got the same results. $P(C \mid +)=.64$ which means $P(C^c \mid +)=.36$

For the next part we have to adjust our sample space. We only consider experiments that provided + results.

This means our $P(C \mid +)=.64$ and $P(C^c \mid +)=.36$ become our new $P(C)$ and $P(C^c)$

Assuming statistical power and p-value remain the same for the second round we can perform the same analysis.

$$\begin{aligned}
 P(C \mid ++) &= \frac{.8 \cdot .64}{.8 \cdot .64+.05 \cdot .36} \\
 &=.9715
 \end{aligned}$$

This seems like a pretty reasonable conclusion. The problem with this analysis is that it assumes test results can actually be modeled by use of a random variable. In the real world, confounding variables like sampling methods, experiment design, researcher cognitive biases, etc, end up being replicated. Close reproduction studies are most effective for identifying data dredging (p-hacking) and statistical anomalies while reproduction with adjustment to methods can identify poorly designed studies. The unfortunate result of this wonderful mess is a popular cultural distrust in science, even when scientific consensus exists (vaccines, climate change, evolution).