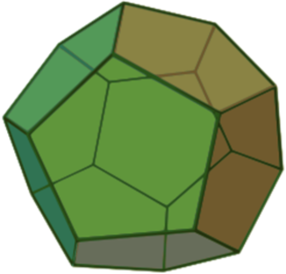
Unit 6: Probability

Expected values

Ex 1: Flip a coin 10 times, paying \$1 to play each time. You win \$.50 (plus your \$1) if you get a head. How much should you expect to win?

Ex 2: Roll two dodecahedral (12-sided) dice. You win \$10 (plus your payment to play) if you get doubles. How much should you pay to play for a fair game?



Two similar examples:

From text: **Paradox of the Chevalier de la Méré:** P(at least 1 ace in 4 rolls of die) $>$ P(at least 1 double-ace in 24 rolls of 2 dice)

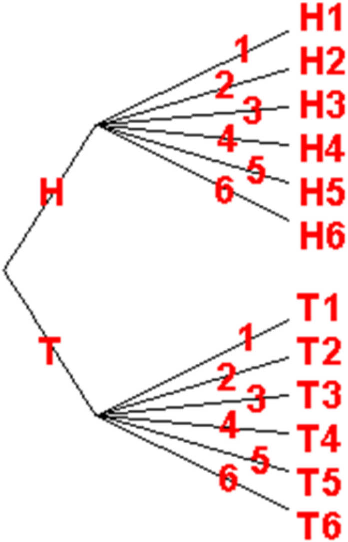
Birthday problem: With 30 people in a room, how likely is it that at least two have the same birth date?

The Birthday problem

# people	P(no match)	P(match)
2	0.99726027	0.00273973
3	0.99179583	0.00820417
4	0.98364409	0.01635591
5	0.97286443	0.02713557
6	0.95953752	0.04046248
7	0.9437643	0.0562357
8	0.92566471	0.07433529
9	0.90537617	0.09462383
10	0.88305182	0.11694818
11	0.85885862	0.14114138
12	0.83297521	0.16702479
13	0.80558972	0.19441028
14	0.77689749	0.22310251
15	0.74709868	0.25290132
16	0.71639599	0.28360401
17	0.68499233	0.31500767
18	0.65308858	0.34691142
19	0.62088147	0.37911853
20	0.58856162	0.41143838
21	0.55631166	0.44368834
22	0.52430469	0.47569531
23	0.49270277	0.50729723
24	0.46165574	0.53834426
25	0.4313003	0.5686997
26	0.40175918	0.59824082
27	0.37314072	0.62685928
28	0.34553853	0.65446147
29	0.31903146	0.68096854
30	0.29368376	0.70631624
31	0.26954537	0.73045463
32	0.24665247	0.75334753
33	0.22502815	0.77497185
34	0.20468314	0.79531686

Tree diagram

Flip a coin, then roll a die, list all alternatives



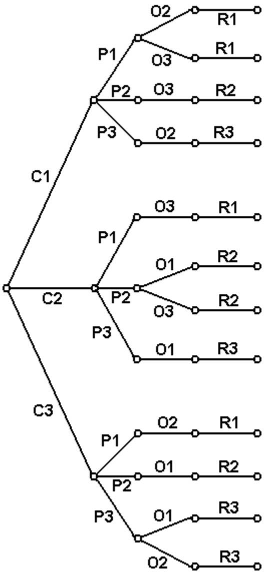
The Monty Hall Problem

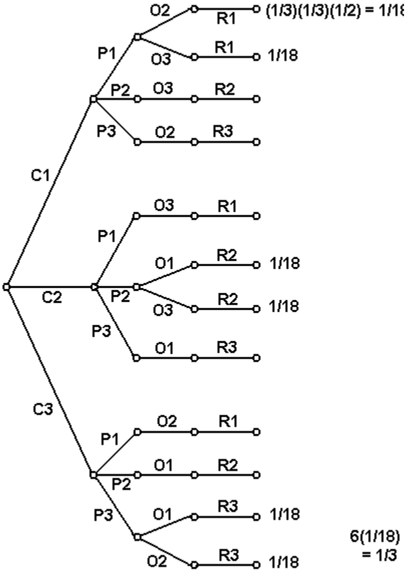
(From Marilyn vos Savant's column) Game show: Three doors hide a car and 2 goats. Contestant picks a door. Host opens one of the other doors to reveal a goat. Contestant then may switch to the other unopened door. Is it better to stay with the original choice or to switch; or doesn't it matter?

Marilyn's answer: Switch!

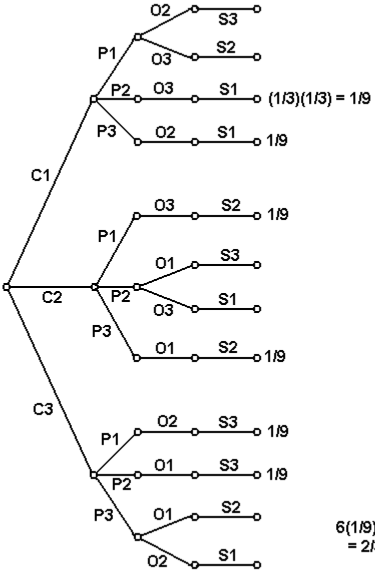
Many respondents: Doesn't matter. ("You're the goat!")

Tree diagram of Stayer's possible games





6(1/18)
= 1/3



6(1/9)
= 2/3

Math stuff about binomial coefficients

They're called that because they are the coefficients of x and y in the expansion of $(x+y)^n$:

$$C(n,0)x^n + C(n,1)x^{n-1}y + C(n,2)x^{n-2}y^2 \\ + \cdots + C(n,n-1)xy^{n-1} + C(n,n)y^n$$

For small n, compute $C(n,k)$ with "Pascal's triangle": 1's in first row and column, then each entry is sum of the two diagonally above.

1
1 1
1 2 1
1 3 3 1
1 4 6 4 1
1 5 10 10 5 1
1 6 15 20 15 6 1

(More from Marilyn vos Savant's column)
Suppose we assume that 5% of the people are drug users. A drug test is 95% accurate (i.e., it produces a correct result 95% of the time, whether the person is using drugs or not). A randomly chosen person tests positive. Is the person highly to be a drug user?

Marilyn's answer: Given your conditions, once the person has tested positive, you may as well flip a coin to determine whether she or he is a drug user. The chances are only 50-50. But the assumptions, the make-up of the test group and the true accuracy of the tests themselves are additional considerations.

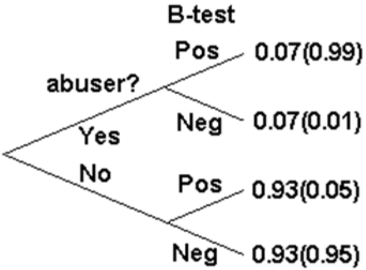
(To see this, suppose the population is 10,000 people; compare numbers of false positives and true positives.)

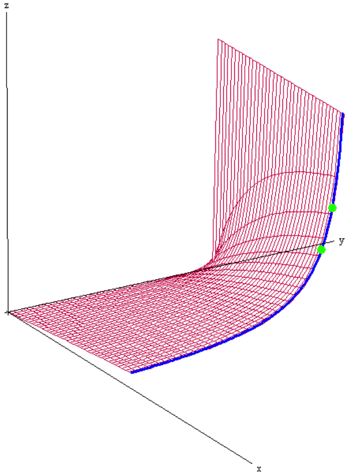
Drug [disease] testing probabilities

Drug [disease] present?	Test positive	Test negative	Sum
Yes	"Sensitivity"	False negative	1
No	False positive	"Specificity"	1

	pos	neg
abuser	.99	.01
clean	.05	.95

Ex: Suppose the Bovine test for lactose abuse has a sensitivity of 0.99 and a specificity of 0.95; and that 7% of a certain population abuses lactose. If a person tests positive on the Bovine test, how likely is it that (s)he really abuses lactose?



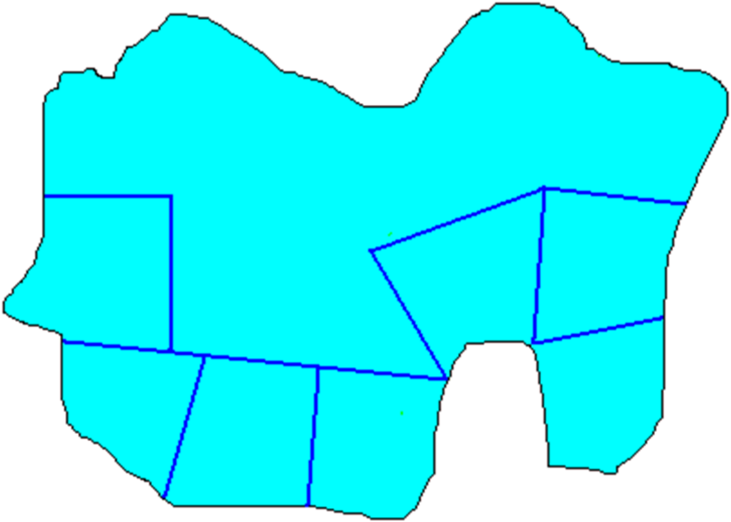


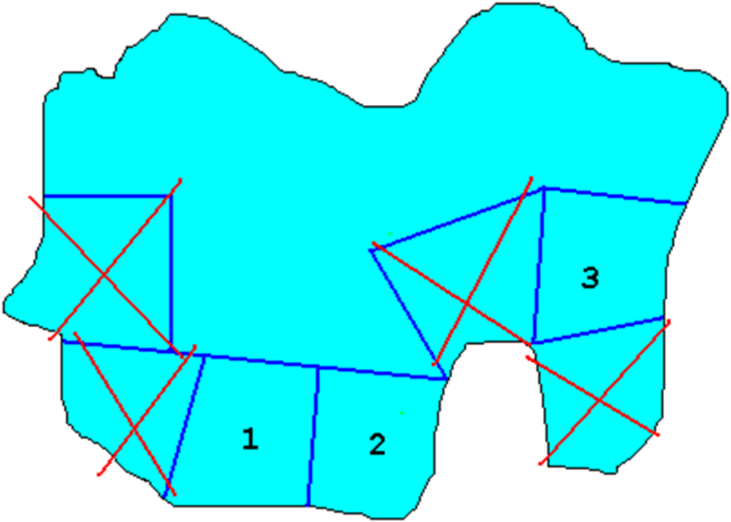
Assuming 7% of population is really positive:
x = sensitivity
y = specificity
$z = P(\text{pos test} \implies \text{pos})$

curve: $x = .99$
points: $(.99, .95, .6)$ and $(.99, .90, .42)$

Counting dragonflies

(thanks to Profs. V. MacMillen and R. Arnold)




1
2
3

Only two pairs

30 censuses altogether, 17 with only two pairs.

Of those 17, 12 had both in same plot.

Do they prefer to lay eggs in proximity?

Censuses with > 2 pairs

P1	P2	P3
0	0	3
0	3	0
3	0	0
1	0	2
1	2	0
3	0	0
1	2	0
2	1	0
1	2	0
0	0	3
1	2	0

P1	P2	P3
4	0	0
0	0	4
3	1	0
4	0	0
0	4	0
0	2	2

Up to 12 at the same time

P1	P2	P3
3	2	0
0	0	5
0	3	2
3	1	1
2	3	0
2	3	0
1	3	1
1	0	4
1	2	2
0	1	4
0	0	5

With 3 pairs