Prob/Stat&*Mathematica*

Authors: Bruce Carpenter, Bill Davis, Michael Raschke and Jerry Uhl

Producer: Bruce Carpenter

Publisher: Math Everywhere, Inc.

Prob.03.3 Probabilities
GIVE IT A TRY!

Mathematica Initializations

MCB 11/5/07: Thanks, your explanation did clear it up (see below). I had expected the linear algebra course to be harder than the probability theory course, but the reverse seems to be true. The linear algebra course builds on concepts that are familiar from vector calculus, but the prob. course involves new concepts and also a new way of thinking. Stretching those mental muscles...

RC: 11/5/07: For you and I both.
Linear Algebra is usually harder for most people, but you have done very well with the geometrical aspects of vector calculus.
There are a few hard things coming in Linear Algebra.

G.1) Calculus Cal bumps into *Mathematica* Fats

G.2) The indicator function, unions and intersections

RC: 11/2/07: So I had to think about this stuff for more than a few hours, as it does boggle the mind a bit on how it all works.
See if my explanation below makes sense. I can make a movie discussion about it if need be.

G.3) Probability calculations in context:

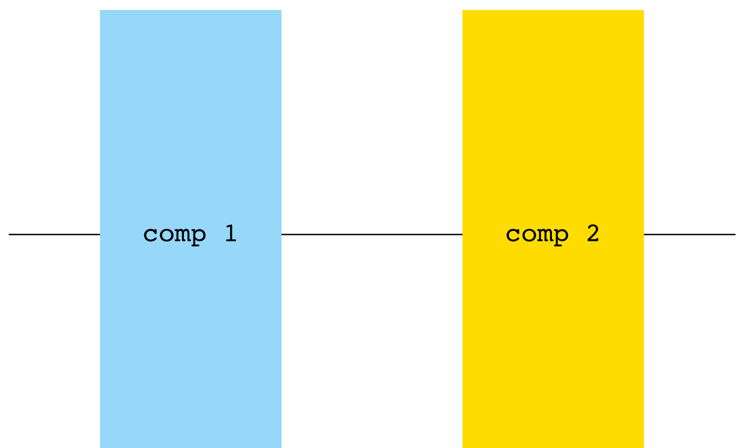
Series wiring versus parallel wiring

This problem was adapted from Jim Pitman's book, "Probability" (Springer Texts in Statistics, Springer-Verlag New York, 1993).

G.3.a.i) Two independent components in series

A system is built using two components wired together in series:

```
wire = Graphics[Line[{{-4, 0}, {4, 0}}]];
comp1 = Graphics[{LightSkyBlue, Polygon[{{-3, -1}, {-1, -1}, {-1, 1}, {-3, 1}}]}];
comp2 = Graphics[{Gold, Polygon[{{3, -1}, {1, -1}, {1, 1}, {3, 1}}]}];
labels = {Graphics[Text["comp 1", {-2, 0}]], Graphics[Text["comp 2", {2, 0}]]};
Show[wire, comp1, comp2, labels];
```



Component 1 has probability 0.93 of working for one year.
Component 2 has probability 0.87 of working for one year.

Assuming that failures of the components are **independent**, estimate the probability that the system composed of both components wired in series will work for one year.

Since the components are wired in series, the system will fail if one or both of the components fails. The probability that both components will work for one year is Prob[c1works $\cap$ c2works], and since the failures of c1 and c2 are independent, the calculation is simply

Prob[c1works ⋂ c2works] = Prob[c1works] Prob[c2works].

Note that the probability of both c1 and c2 working for one year is the same as the probability that neither c1 nor c2 fail during the year.

```
Clear[prob];
prob[c1works] = 0.93;
prob[c2works] = 0.87;
prob[bothwork] = prob[c1works] prob[c2works]
```

```
0.8091
```

The probability that the system composed of both components wired in series will work for one year is 0.8091.

RC: 11/2/07: Good. Doing this computation in reverse, using failures:
Prob[F1 ⋃ F2] = one of the components fails
prob[F1 ⋂ F2] = both of the components fail
Prob[F1 ⋃ F2] = prob[F1] + prob[F2] – prob[F1 ⋂ F2]
Independence means:
prob[F1 ⋂ F2] = prob[F1] * prob[F2]
Prob[F1 ⋃ F2] = 0.07 + 0.13 - (0.07*0.13) = 0.1909
Therefore, flipping it around,
Prob[W1 ⋂ W2] = 1 - Prob[F1 ⋃ F2] = 1 - 0.1909 = 0.8091

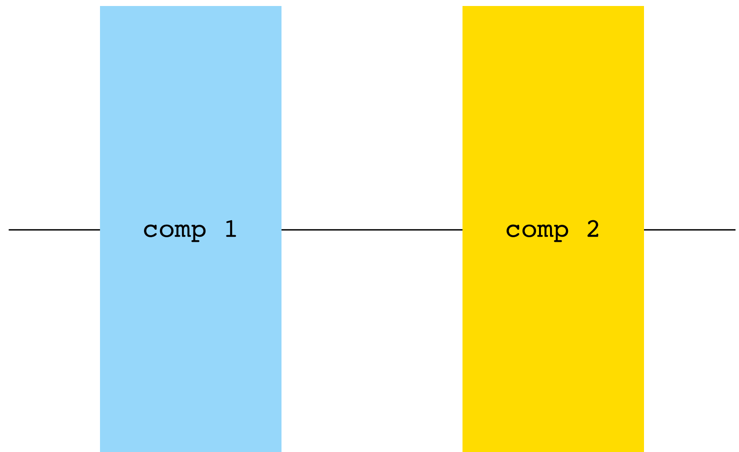
```
1 - 0.1909
```

```
0.8091
```

□ G.3.a.ii) Two not independent components in series

A system is built using two components wired together in series:

```
wire = Graphics[Line[{{-4, 0}, {4, 0}}]];
comp1 = Graphics[{LightSkyBlue, Polygon[{{-3, -1}, {-1, -1}, {-1, 1}, {-3, 1}}]}];
comp2 = Graphics[{Gold, Polygon[{{3, -1}, {1, -1}, {1, 1}, {3, 1}}]}];
labels = {Graphics[Text["comp 1", {-2, 0}]], Graphics[Text["comp 2", {2, 0}]]};
Show[wire, comp1, comp2, labels];
```



Component 1 has probability 0.93 of working for one year.
Component 2 has probability 0.87 of working for one year.

Assuming that:

Prob[Component 2 fails, Given that Component1 fails] = p,

estimate the probability that the system composed of both components wired in series will work for one year.

Again, if at least one component fails, the system fails.
The probability that at least one component fails is
Prob[c1fails $\bigcup$ c2fails] = Prob[F1 $\bigcup$ F2] and the probability that both components work is

Prob[bothwork] = Prob[W1 $\bigcap$ W2] = 1 − Prob[F1 $\bigcup$ F2]

We know that

Prob[F1 $\bigcup$ F2] = F1 + F2 − F1 $\bigcap$ F2

Because failure of the components is not independent,

prob[F1 $\bigcap$ F2] $\neq$ prob[F1] prob[F2].

Instead, using the conditional probability formula, you have

prob[F1 $\bigcap$ F2] = prob[F1] prob[F2, givenF1]

So

Prob[W1 $\bigcap$ W2] = 1 − (F1 + F2 − F1 $\bigcap$ F2)
Prob[W1 $\bigcap$ W2] = 1 − F1 − F2 + F1 $\bigcap$ F2
Prob[W1 $\bigcap$ W2] = 1 − F1 − F2 + (prob[F1] prob[F2, givenF1])

```
Clear[prob, bothwork];
prob[F1] = 1 - 0.93
prob[F2] = 1 - 0.87
prob[F2, givenF1] = p

prob[bothwork] = 1 - ( prob[F1] + prob[F2] - (prob[F1] prob[F2, givenF1]) )
```

```
0.07
0.13
p
0.8 + 0.07 p
```

The probability that the system composed of both components wired in series will work for one year if failure of the components is not independent is $0.8 + 0.07\,p$.

```
0.8 + 0.07 p /. p → 0.5
0.8 + 0.07 p /. p → 1
```

```
0.835
0.87
```

?? That can't be right. Then the probability of the system working for one year would be *higher* than it was in G.3.a.i) with the assumption of independence. Where did I go wrong?

```
0.07 * 0.13
```

```
0.0091
```

RC: 11/2/07: Here are my computations
Prob[F1 $\bigcup$ F2] = one of the components fails
prob[F1 $\bigcap$ F2] = both of the components fail
Prob[F1 $\bigcup$ F2] = prob[F1] + prob[F2] – prob[F1 $\bigcap$ F2]
Now this formula says: - prob[F1 $\bigcap$ F2] means: take away the times that both might be failing, i.e. we double-count the failing.
if independent, then prob[F1 $\bigcap$ F2] = (0.07*0.13) ... a low probability that both fail simultaneously

non-Independence means: Conditional Probability Formula:

$$\text{Prob}[A, X\text{given}Y] = \frac{\text{Prob}[A\bigcap Y, X]}{\text{Prob}[Y, X]}$$

Adapted to this problem: Y = {0} (failure), A = F1, X = F2

Prob[F1, F2 fails] * Prob[F1] = Prob[F1⋂F2]

prob[F1 ⋂ F2] = prob[F1] prob[F2,givenF1] = 0.07 * p = potentially larger number = higher probability of failure of both.
(p = 0.13 is magic cutoff; if p < 0.13, then simultaneous failure is less likely than independent;
if p > 0.13, then simultaneous failure is more likely than independent)

For example, if p=1, this means that Prob[F1, F2 fails] = 1, so that if F2 fails, F1 will certainly fail;
i.e. F1 is actually dependent upon F2. i.e. F2 is at a power source, and if it dies, F1 dies.
So then the total contribution of "we double-counted the failing" is:
0.07*1 = 0.07 = the prob[F1].
This makes sense since if F2 fails, F1 will fail, so the probability of 0.07 for F1 failing will always
get counted twice.

For example, if p=0.5, this means that Prob[F1, F2 fails] = 0.5, so that if F2 fails, F1 will fail 50% of the time.
i.e. F1 is dependent upon F2 1/2 the time..
So then the total contribution of "we double-counted the failing" is:
0.07*0.5 = 0.035 = 1/2 * the prob[F1].
This makes sense since if F2 fails, F1 will fail 1/2 of the time, so the probability of 0.07 for F1 failing will
get counted twice 1/2 of the time.

"Independence" is measured when the dependency of F1 on F2 is EXACTLY THE SAME as if there
was independence, i.e. Prob[F1, F2 fails] = 0.13. In our minds, there is still dependency. But from a probability
point of view, the dependency is the same measurement as if independent, i.e 0.07 * 0.13. This is what B4 and B5
are all talking about - measuring "independence". In a physics environment, it might be difficult to determine
if two components are independent or not - the term "independent" is

applied if you can measure their "counting
twice" to be exactly 0.07 * 0.13, i.e. Prob[F1, F2 fails] = Prob[F2].

Prob[F1 ⋃ F2] = 0.07 + 0.13 - (0.07p) = 0.20 - 0.07p
Therefore, flipping it around,
Prob[W1 ⋂ W2] = 1 - Prob[F1 ⋃ F2] = 1 - (0.20 - 0.07p) = 0.80 + 0.07p
So your formula is correct. Let's try to talk our way through it:

A series circuit only needs 1 to fail for the circuit to fail. If p=1, then F1 is dependent upon F2.
Therefore, the working should be Prob[W2] = 0.87.
Does our formula match here? 0.80 + 0.07*1 = 0.87.
What about p=0.5?
F1 will fail 50% of the time when F2 fails.
So failing should be: 0.07 (F1) + 0.13 (F2) - 0.07(0.5) (counting twice those dependent failures, which are high) = 0.165
So working = 1 - 0.165 = 0.835
Under independence, working is 80.91%. So the circuit being dependent at 50% means the probability of working
is higher - 83.5% - than for 2 independent circuits. This is because the independent failure of F1 is so low.

When we have a dependent condition, like p=0.5,
it is not like F1 fails more.
F1 fails at 7% and that's it.
The question is: how much of the 7% is due to F2 failing? Because we should not count
these "both failed" situations twice. If p=0.5, then 0.07*p is the amount of F1 failing because of F2 failing ---
3.5% of F1's failing was on its own, but 3.5% of F1's failing was because of F2, so for a series circuit,
we should not count this second 3.5% twice. In this way, one can think of the calculation at:
3.5% F1 failure + 13% F2 failure = 16.5% failure --> working 1- 0.165 = 83.5%

Yes, better than the independent probability.   F2 will fail only 13% of the time, and
F1 will fail - well, 7% of the time, but 50% of that failure is due to F2 failing, so
it really only will fail 3.5% of the time (without F2 failing).

```
0.05 + 0.2 - 0.05 * 0.2
1 - (0.165)
```

```
0.24
0.835
```

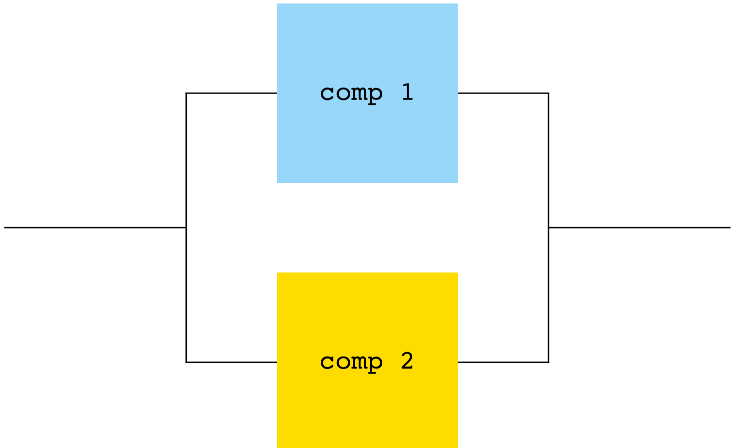
MCB 11/5/07: Thanks, that does clear it up. What confused me was not seeing that the dependent failure was *part of the total failure rate for the dependent component*, not in addition to it. When I first approached the problem I was trying to break it down by separating prob[F2, givenF1] from prob[F2, NOT givenF1] but just got muddled. Now it makes sense.

RC: 11/5/07: Took me a few minutes (like a few hours) to sort it out in my head, too.

G.3.b.i) Two independent components in parallel

A system is built using two components wired together in parallel:

```
wire = {Graphics[Line[{{0, 0}, {4, 0}}]], Graphics[Line[{{0, 3}, {4, 3}}]],
 Graphics[Line[{{0, 0}, {0, 3}}]], Graphics[Line[{{4, 0}, {4, 3}}]],
Graphics[Line[{{-2, 1.5}, {0, 1.5}}]], Graphics[Line[{{4, 1.5}, {6, 1.5}}]]};
comp1 = Graphics[{LightSkyBlue, Polygon[{{3, 2}, {1, 2}, {1, 4}, {3, 4}}]}];
comp2 = Graphics[{Gold, Polygon[{{3, -1}, {1, -1}, {1, 1}, {3, 1}}]}];
labels = {Graphics[Text["comp 1", {2, 3}]], Graphics[Text["comp 2", {2, 0}]]};
Show[wire, comp1, comp2, labels];
```



Component 1 has probability 0.93 of working for one year.
Component 2 has probability 0.91 of working for one year.

Assuming that failures of the components are independent, estimate the probability that the system composed of both components wired in parallel will work for one year.

Because the components are wired in parallel, the system will work if at least one component works.

Prob[at least one works] = Prob[c1works or c2works or c1andc2work]

Prob[at least one works] = prob[W1 $\bigcup$ W2]

Prob[at least one works] =

prob[W1] + prob[W2] – prob[W1 $\bigcap$ W2] (because there is overlap)

If failures of the components are independent, the probability of no failures (prob[W1 $\bigcap$ W2]) is

prob[W1 $\bigcap$ W2] = prob[W1] prob[W2]

So prob[W1 $\bigcup$ W2] = prob[W1] + prob[W2] – prob[W1] prob[W2]

```
Clear[prob];
prob[W1] = 0.93;
prob[W2] = 0.91;
prob[W1unionW2] = prob[W1] + prob[W2] - (prob[W1] prob[W2])
```

```
0.9937
```

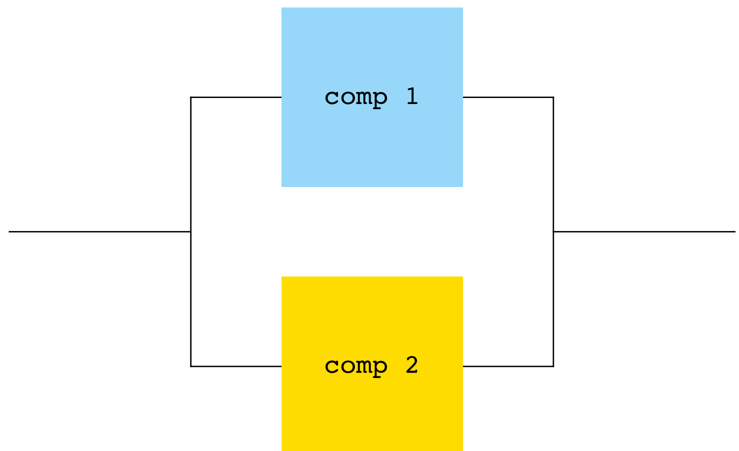
The probability that the system composed of both components wired in parallel will work for one year is 0.9937.

RC: 11/5/07: Parallel is must easier: both must fail.

G.3.b.ii) Two non-independent components in parallel

A system is built using two components wired together in parallel:

```
wire = {Graphics[Line[{{0, 0}, {4, 0}}]], Graphics[Line[{{0, 3}, {4, 3}}]],
 Graphics[Line[{{0, 0}, {0, 3}}]], Graphics[Line[{{4, 0}, {4, 3}}]],
Graphics[Line[{{-2, 1.5}, {0, 1.5}}]], Graphics[Line[{{4, 1.5}, {6, 1.5}}]]};
comp1 = Graphics[{LightSkyBlue, Polygon[{{3, 2}, {1, 2}, {1, 4}, {3, 4}}]}];
comp2 = Graphics[{Gold, Polygon[{{3, -1}, {1, -1}, {1, 1}, {3, 1}}]}];
labels = {Graphics[Text["comp 1", {2, 3}]], Graphics[Text["comp 2", {2, 0}]]};
Show[wire, comp1, comp2, labels];
```



Component 1 has probability 0.93 of working for one year.
Component 2 has probability 0.91 of working for one year.

Assuming that:

Prob[Component2 fails, Given that Component1 fails] = 0.55,

estimate the probability that the system composed of both components wired in parallel will work for one year.

Because the components are wired in parallel, the system will work if at least one component works. This is the same as saying that the system will only fail if both components fail.

Prob[both c1 and c2 fail] = prob[F1 $\cap$ F2]

Prob[at least one component works] = 1 – prob[F1 ⋂ F2]

Because failure of the components is not independent,

prob[F1 ⋂ F2] ≠ prob[F1] prob[F2].

Instead, using the conditional probability formula, you have

prob[F1 ⋂ F2] = prob[F1] prob[F2, givenF1]

So 1 – prob[F1 ⋂ F2] = 1 – (prob[F1] prob[F2, givenF1])

```
Clear[prob];
prob[F1] = 1 - 0.93
prob[F2] = 1 - 0.91
prob[F2, givenF1] = 0.55;
prob[F1intersectF2] = 1 - (prob[F1] prob[F2, givenF1])
```

```
0.07
0.09
0.9615
```

The probability that the system composed of both components wired in parallel will work for one year if component failure is not independent is 0.9615. Note that for components wired in parallel, dependence *decreases* the probability that the system will work for one year by *increasing* the probability that both components will fail.

```
Clear[F1, F2];
{F1, F2} = {0.07, 0.09};
F1 * F2
F2givenF1 = 0.55;
F1 * F2givenF1
```

```
0.0063
0.0385
```

The possible extremes here are:

- c2 always fails if c1 fails (complete dependence)

  Then prob[F2, givenF1] = 1

  Probability that both c1 and c2 fail is

F1 ⋂ F2 = F1 (F2, givenF1) = F1 (1) = F1

F1 ⋂ F2 = 0.07

- c2 fails at the same rate regardless of whether c1 fails (complete independence)

Then prob[F2, givenF1] = prob[F2, not given F1] = F2

Probability that both c1 and c2 fail is

$F1 \bigcap F2 = F1\ (F2, \text{givenF1}) = F1\ (F2, \text{not given } F1) = F1 * F2$

$F1 \bigcap F2 = 0.07 * 0.09 = 0.0063$

Note that prob[F2, givenF1] *can not be lower* than

prob[F2, independent of F1] = F2. But if there is any dependence at all,

prob[F2, givenF1] > prob[F2, independent of F1] = F2

so $F1 \bigcap F2 = F1\ (F2, \text{givenF1}) >$

$F1 \bigcap F2 = F1\ (F2, \text{independent of } F1) = F1 * F2$

and $F1\ (F2, \text{independent of } F1) \leq F1 \bigcap F2 \leq F1\ (F2, \text{givenF1})$

Also note that prob[F2, givenF1] includes *both* the probability that c2 fails because c1 fails, and the probability that c2 fails independently of c1. In other words,

prob[F2, givenF1] = prob[F2, independent of F1] + prob[F2, caused by F1]. So if you know the probability that c2 fails by itself (independently of c1), you can calculate the probability that c1 causes c2 to fail:

prob[F2, caused by F1] = prob[F2, givenF1] – prob[F2, independent of F1]. Even in the extreme case that prob[F2, givenF1] = 1 (complete dependence), it is very unlikely that 100% of c2 failures are caused by c1 failing; c1 must fail on its own sometimes, even if the probability is very low.

In the current example,

prob[F2, independent of F1] = F2 = 0.09

prob[F2, givenF1] = 0.55

prob[F2, caused by F1] = 0.55 − 0.09 = 0.46

RC: 11/5/07: Looks good to me. "Independence" to probability people is p = prob[F2] / prob[F1]. Not real independence, but just "in terms of the numbers, it is just the same computation as independence". Of course, to a logician, this is a painful use of language.

G.4) Dicey calculations

G.5) Probability calculations in context:

Tattoos and tongue rings

G.6) Probability calculations in context:

False positives in breast cancer tests

G.7) Independence

G.8) Probability calculations in context:

Probability and aces

G.9) Probability calculations in context:

Info you can use when you go to Vegas and hit the craps table

G.10) Probability calculations in context:

Gambler's ruin: Some things every serious gambler should know

G.11) Probability calculations in context:

The birthday problem

G.12) Randall Swift's eighteen sided computer dice

G.13) Actuarial exam probability problems from the Society of Actuaries