Lecture 2

David Aldous

28 August 2015

The specific examples I'm discussing are not so important; the point of these first lectures is to illustrate a few of the 100 ideas from STAT134.

Ideas used in Lecture 1.

- Bayes rule.
- $\mathbb{E}g(X) = \sum_x g(x)\mathbb{P}(X = x)$.
- Fair bet means $\mathbb{E}(\mathrm{gain}) = 0$.
- Indicator r.v.'s $\mathbb{1}(A)$ useful notation.
- Law of large numbers.

4. Fair price for option on Normal payoff.

Consider Z with Normal$(0, 1)$ distribution.
The fair price today to receive Z tomorrow $= \mathbb{E}Z = 0$.

What is fair price for an "option" to receive Z tomorrow for cost c?
That is, a contract that says
tomorrow, after you see the value of Z, you can buy it for c if you wish.

If you buy the option, then tomorrow

if $Z > c$ then you have payoff $Z - c$
if $Z < c$ then you have payoff 0.

In other words, your payoff is $\max(Z - c, 0)$, so fair price is

$$\mathbb{E} \max(Z - c, 0).$$

Need to calculate $\mathbb{E}\max(Z-c,0)$, where Z has the standard Normal density function $\phi(z)=(2\pi)^{-1/2}\exp(-z^2/2)$. Recall general formula

$$\mathbb{E}g(Z)=\int g(z)\phi(z)\,dz.$$

So here

$$\mathbb{E}\max(Z-c,0)=\int_c^\infty (z-c)\phi(z)dz.$$

Because $\frac{d}{dz}\phi(z)=-z\phi(z)$ we can integrate by parts to get

$$=\phi(c)-c\bar{\Phi}(c)$$

for

$$\bar{\Phi}(c)=\int_c^\infty \phi(z)dz.$$

5. Decreasing dice rolls.

Throw a die 3 times – what is the chance that the successive numbers are strictly decreasing, like $5, 3, 2$*?*

Can do by counting – there are 20 possibilities, so chance $= 20/6^3$. But how to do the general question:

Throw a hypothetical m-sided (numbers $1, 2, \ldots, m$*) die k times, for* $k \le m$*. What is the chance* $p(k, m)$ *that the successive numbers are strictly decreasing?*

Now we need to get more organized . . . . . .

Throw a hypothetical m-sided (numbers $1, 2, \ldots, m$) die k times.

$$p(k, m) = \mathbb{P}(\text{successive numbers are strictly decreasing}).$$

- -

Small trick: for this to happen, the numbers must be all different.
If the k numbers are all different, then each of the $k!$ possible orders are equally likely ("argument by symmetry"), so the conditional probability of being strictly decreasing $= 1/k!$ So

$$p(k, m) = \frac{1}{k!} \times \mathbb{P}(\ k \text{ numbers all different }\ )$$

But we know from the "birthday problem" that

$$\mathbb{P}(\ k \text{ numbers all different }\ ) = \frac{m-1}{m} \times \frac{m-2}{m} \times \ldots \times \frac{m-k+1}{m} = \frac{m!}{m^k (m-k)!}$$

and we find

$$p(k, m) = \frac{1}{m^k} \times \binom{m}{k}$$

which suggests another way to derive this answer.

6. The 3rd formula for variance.

$$\mathbb{E}X = \sum_x x\mathbb{P}(X = x) \text{ or } \int xf(x)dx.$$

Writing $\mu = \mathbb{E}X$, the definition of variance is

$$\mathrm{var}\ X = \mathbb{E}(X - \mu)^2$$

and a line of algebra gives an equivalent formula

$$\mathrm{var}\ X = EX^2 - \mu^2.$$

Recall $\mathrm{s.d.}(X) = \sqrt{\mathrm{var}\ X}$ is the "interpretable" measure of spread of the RV X, which scales in the natural way:

$$\mathrm{s.d.}(cX) = |c| \times \mathrm{s.d.}(X); \quad \mathrm{var}\ (cX) = c^2 \mathrm{var}\ (X).$$

Variance is mathematically convenient because of the property:

$$\text{if } X, Y \text{ independent then } \mathrm{var}\ (X+Y) = \mathrm{var}\ (X) + \mathrm{var}\ (Y).$$

But note this implies

$$\text{if } X, Y \text{ independent then } \mathrm{var}\ (X-Y) = \mathrm{var}\ (X) + \mathrm{var}\ (Y)$$

with a "plus" not a "minus". This leads to the " 3rd formula for variance":

$$\mathrm{var}\ X = \tfrac{1}{2}\mathbb{E}(X_1 - X_2)^2$$

where X_1 and X_2 are independent r.v.'s distributed as X. This formula has an intuitive "variability of realizations" interpretation.

7. Inventing extra structure in a problem.

This sounds like cheating, but we are not "making extra assumptions", but instead we are merely defining extra random variables to help analyze the given random variables.

Here's an example from STAT134.

7a: best-out-of-$(2k-1)$ **contest.** [e.g. baseball World Series]

Teams A and B play until one team has won k games. The model is that $\mathbb{P}(A \text{ beats } B) = p$, independently for each game. The number of games played, G, is random with possible values $k \le G \le 2k-1$.

Problem: *Calculate* $\mathbb{P}(A \textit{ wins series })$.

Solution. Imagine they play all $2k-1$ games, so A wins some number of all the games, say Y. The event $\{A \text{ wins series }\}$ is the same as the event $\{Y \geq k\}$. So

$$\mathbb{P}(A \text{ wins series }) = \mathbb{P}(Y \geq k)$$

and Y has Binomial$(2k-1, p)$ distribution.

7b. Put $k \geq 3$ points at random (independent uniform) on the circumference of a circle. These points are the vertices of a polygon; what is the probability $p(k)$ of the event

$$G = \{ \text{ the polygon does not contain the center of the circle} \}?$$

There is a simple argument based on a trick that you or I would never think of. Implement "k random points" in 2 stages.

Stage 1. Create k pairs of diametrically-opposite points.
Stage 2. Pick one point from each pair.

Stage 1. Create k pairs of diametrically-opposite points.
Stage 2. Pick one point from each pair.

Condition on the result of Stage 1. What is the conditional probability that, after the Stage 2 process, the event A occurs?

By drawing a picture, the event G occurs if and only if, in Stage 2, we pick k adjacent points. There are $2k$ possible sets of such points, so

$$\mathbb{P}(G | \text{ result of Stage 1 }) = 2k/2^k.$$

Because this is the same whatever the result of Stage 1, we have

$$\mathbb{P}(G) = 2k/2^k = k/2^{k-1}.$$

8. Size-biasing

A topic maybe not discussed in STAT134. Suppose each child is in some family; we can then consider

- X = number of children in a uniform random family
- $\widetilde{X}$ =number of children in the family of a uniform randomly picked child.

These are different! To see why, write N = number of families. Then

- Number of families with i children = ???
- Number of children in i-child families = ???
- Total number of children = ????

$$\mathbb{P}(\widetilde{X} = i) = ???$$

Say $\widetilde{X}$ has the **size-biased** distribution of X.

- $X =$ number of children in a uniform random family
- $\widetilde{X} =$number of children in the family of a uniform randomly picked child.
- $N =$ number of families.

We calculate

- Number of families with i children $= N\mathbb{P}(X = i)$
- Number of children in i-child families $= i \times N\mathbb{P}(X = i)$
- Total number of children $= \sum_i i \times N\mathbb{P}(X = i) = N\ \mathbb{E}X$

$$\mathbb{P}(\widetilde{X} = i) = \frac{i \times N\mathbb{P}(X = i)}{N\ \mathbb{E}X} = \frac{i\ \mathbb{P}(X = i)}{\mathbb{E}X}.$$

Say $\widetilde{X}$ has the **size-biased** distribution of X.

It's a good exercise to express the distribution of X in terms of the distribution of $\widetilde{X}$, and to give formulas for the expectations in terms of the other distribution.

- $X =$ number of children in a uniform random family
- $\widetilde{X}$ =number of children in the family of a uniform randomly picked child.

$$\mathbb{P}(\widetilde{X} = i) = \frac{i\ \mathbb{P}(X = i)}{\mathbb{E}X}.$$

$$\mathbb{E}\widetilde{X} = \sum_i i\ \mathbb{P}(\widetilde{X} = i) = \frac{\sum_i i^2\ \mathbb{P}(X = i)}{\mathbb{E}X} = \frac{\mathbb{E}X^2}{\mathbb{E}X}.$$

To calculate $\mathbb{E}X$ from the distribution of $\widetilde{X}$, observe

$$\sum_i i^{-1}\mathbb{P}(\widetilde{X} = i) = \frac{\sum_i\ \mathbb{P}(X = i)}{\mathbb{E}X} = \frac{1}{\mathbb{E}X}$$

and so

$$\mathbb{E}X = \frac{1}{\sum_i i^{-1}\mathbb{P}(\widetilde{X} = i)} \qquad \text{(“harmonic mean”)}$$

$$\mathbb{P}(X = i) = \frac{i^{-1}\mathbb{P}(\widetilde{X} = i)}{\sum_j j^{-1}\mathbb{P}(\widetilde{X} = j)}.$$

This effect occurs in other contexts, such as class size.

If a small Department offers two courses, with enrollments 90 and 10, then

average class (faculty viewpoint) $= (90 + 10)/2 = 50$
average class (student viewpoint) $= (90 \times 90 + 10 \times 10)/100 = 82.$