MTH135/STA104: Probability

Homework # 5 Due: Tuesday, Oct 4, 2005

Prof. Robert Wolpert

1. Let X_1 and X_2 be the numbers on two independent rolls of a fair die; set

$$Y_1 \equiv \min(X_1, X_2) \qquad\qquad Y_2 \equiv \max(X_1, X_2)$$

a) Give the joint distribution of X_1 and X_2

The thirty-six possibilities are all equally likely, so $\mathsf{P}[X_1 = i, X_2 = j] = \frac{1}{36}$ for $i, j \in \{1, 2, 3, 4, 5, 6\}$.

b) Give the distribution of Y_1

Find the joint distribution (below) first; then sum the margins. The six possible values of Y_1 have probabilities 11/36, 9/36, 7/36, 5/36, 3/36, 1/36, respectively, or $\mathsf{P}[Y_1 = j] = (13 - 2j)/36$ for $j = 1, ..., 6$.

c) Give the joint distribution of Y_1 and Y_2

In a table,

Y_1 \ Y_2	1	2	3	4	5	6
1	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
2	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
3	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$	$\frac{1}{18}$
4	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{18}$
5	0	0	0	0	$\frac{1}{36}$	$\frac{1}{18}$
6	0	0	0	0	0	$\frac{1}{36}$

d) Give the distribution of $Z \equiv Y_2 - Y_2$. What is its most likely value?

There are six possible values, 0..5, with probabilities $1/18$, $2/18 = 1/9$, $3/18 = 1/6$, $4/18 = 2/9$, $5/18$,and $6/36 = 1/6$, respectively. Evidently the most likely value is $\mathsf{P}[Z = 4] = 5/18 \approx 0.2778$.

2. A fair coin is tossed three times; let X denote the number of heads on the first two tosses, Y the number of heads on the last two tosses.

a) Are X and Y independent? Prove it.

No; for example, $\mathsf{P}[X = 0] = \mathsf{P}[Y = 0] = \frac{1}{4}$, while $\mathsf{P}[X = 0, Y = 0] = \frac{1}{8} \neq (\frac{1}{4})^2$.

b) Give the distribution of $Z = X \cdot Y$.

The four possible values of Z have probabilities $\mathsf{P}[Z = 0] = \frac{3}{8}$, $\mathsf{P}[Z = 1] = \frac{1}{4}$, $\mathsf{P}[Z = 2] = \frac{1}{4}$, $\mathsf{P}[Z = 4] = \frac{1}{8}$.

3. Is it possible for two *independent* random variables X and Y (not necessarily with the same distributions) to satisfy $\mathsf{P}[X < Y] = 1$? Why?

Sure. For example, let $X \in \{0, 1\}$ be equally-likely and let $Y \in \{4, 5\}$ be equally-likely. Even simpler, let $X = 0$ and $Y = 1$ with probability one!

4. For some constant $c > 0$ the random variable X takes the value $X = j$ with probability $c \cdot j$ for $j \in \{1, 2, 3, 4\}$.

a) What is the probability that X is an even number?

Evidently $c = 1/10$, so $\mathsf{P}[X \in \{2, 4\}] = 0.20 + 0.40 = 0.60$.

b) How large would you expect the *sum* to be of n independent observations $X_1,\ldots,X_n$ from this distribution?

On average each observation will be $\mathsf{E}[X_j] = \sum x\,\mathsf{P}[X = x] = 1 \cdot \frac{1}{10} + 2 \cdot \frac{2}{10} + 3 \cdot \frac{3}{10} + 4 \cdot \frac{4}{10} = 30/10 = 3$, so the sum of 100 repetitions should be about $3\,n$.

c) What is the probability that n independent observations $X_1,\ldots,X_n$ are all equal? Simplify as much as possible.

$$\begin{aligned}
\sum_{j=1}^{4} \mathsf{P}[X_i = j \text{ for all } 1 \leq i \leq n] &= (1/10)^n + (2/10)^n + (3/10)^n + (4/10)^n \\
&= (1 + 2^n + 3^n + 4^n)/10^n
\end{aligned}$$

5. (from Prob 13, *p.* 160) A box contains $2n$ balls of n different colors, with 2 of each color. Balls are picked at random from the box with replacement until two balls of the same color have appeared. Let X be the number of draws made.

a) Find the probability distribution for X— *i.e.*, find $\mathsf{P}[X = x]$ for every x (hint: find $\mathsf{P}[X > x]$ first, for every integer x).

$$\mathsf{P}[X > x] \;=\; 1 \text{ for } x \le 1 \text{ while, for integers } x \ge 2,$$

$$\mathsf{P}[X > x] \;=\; \overbrace{\frac{2n}{2n} \times \frac{2n-2}{2n} \times \cdots \times \frac{2n+2-2x}{2n}}^{x \text{ terms}} = \prod_{k=0}^{x-1}(1 - k/n)$$

Thus, for $x \in \mathbb{N}$,

$$\mathsf{P}[X = x] \;=\; \frac{x-1}{n}\prod_{k=0}^{x-2}(1 - k/n)$$

b) Find the limit as $n \to \infty$ of the probability that it takes more than $\sqrt{n}$ draws to find two of the same color, if the box contains n different colors— *i.e.*, find $\lim_{n\to\infty} \mathsf{P}[X > \sqrt{n}]$ (Hint: use an exponential approximation for $\mathsf{P}[X > x]$).

$$\mathsf{P}[X > x] = \prod_{k=0}^{x-1}(1 - k/n) \approx \prod_{k=0}^{x-1} e^{-k/n} = e^{-\frac{1}{n}\sum_{k<x} k} \approx e^{-x^2/2n},$$

so $\mathsf{P}[X > \sqrt{n}] \approx e^{-1/2} \approx 0.6065$.

6. Let X_1, X_2,..., X_n be a sequence of random variables. If each pair (X_i, X_j) are independent for $1 \le i < j \le n$, does it follow that $\{X_1, X_2, ..., X_n\}$ are indepndent? Sketch a proof or counterexample.

No; toss two fair coins and let X_1 be the number of heads on the first toss, X_2 the number of heads on the second toss, and let X_3 be one if $X_1 = X_2$ and zero otherwise.

7. Let X be drawn uniformly from the interval $[0, 1]$ and let Y be selected uniformly from the set $\{0, 1, 2\}$, with X and Y independent. Set $Z = X + Y$. Find the function $F(z) = \mathsf{P}[Z \le z]$ for every number $-\infty < z < \infty$. (Hint: Consider the different possible values of Y separately)

8. Suppose $\mathsf{E}[X^2] = 5$, $\mathsf{E}[Y^2] = 10$, and $\mathsf{E}[X \cdot Y] = 6$.

a) Find $\mathsf{E}[(X - Y)^2]$.

$$\mathsf{E}[(X - Y)^2] = \mathsf{E}[X^2] - 2\mathsf{E}[X \cdot Y] + \mathsf{E}[Y^2] = 5 - 12 + 10 = 3$$

b) Find the number $t \in \mathbb{R}$ that minimizes $f(t) = \mathsf{E}[(X - t\,Y)^2]$.

$$f(t)\mathsf{E}[(X - t\,Y)^2] = \mathsf{E}[X^2] - 2\,t\,\mathsf{E}[X \cdot Y] + t^2\mathsf{E}[Y^2] = 5 - 12\,t + 10t^2,$$

a quadratic function with minimum where $0 = f'(t) = 20\,t - 12$, or $t = 0.60$.

9. A building has 10 floors above the basement. If 12 people get onto an elevator and if each picks a floor at random from $\{1, 2, ..., 10\}$ to get out, independently of the others,

a) At how many floors do you expect the elevator to stop?

Let $I_k = 1$ if the elevator stops at floor k for $1 \le k \le 10$; the expectation of I_k is just the probability that at least one of the twelve people stop there, one minus the probability that none stop, $\mathsf{E}[I_k] = 1-(9/10)^{12} \approx 0.71757$. The number of floors at which the elevator stops is $S = \sum_{k=1}^{10} I_k$, with expectation $\mathsf{E}[S] = 10 \times [1 - (9/10)^{12}] \approx 7.176$.

b) Let H be the highest floor the elevator reaches. Find the median m for H, *i.e.*, the number such that $\mathsf{P}[H < m] \le \frac{1}{2} \le \mathsf{P}[H \le m]$.

The probability that $[H \le h]$ is $(h/10)^{12}$, so the median will be the smallest $h \in \{1, ..., 10\}$ with $(h/10)^{12} \ge \frac{1}{2}$, *i.e.*, $h \ge 10(1/2)^{1/12} = 9.439$, so $m = 10$. The probability the top floor is reached is $1 - (9/10)^{12} = 0.7176$, well above 50%.

10. Pitman's problem 20 on *p.* 184 asks you to show that the distribution of any random variable X taking three values (he chooses $\{0, 1, 2\}$) is determined completely by the two "moments" $\mu_1 = \mathsf{E}[X]$ and $\mu_2 = \mathsf{E}[X^2]$, by finding formulas for $\mathsf{P}[X = x]$ in terms of μ_1 and μ_2. Show that this would *not* be true for random variables taking (at least) four different values, by finding two random variables X and Y with the *same* first two moments but *different* distributions. Suggestion: Choose $\{\pm 1, \pm 2\}$ for your points and find distinct distributions with $\mu_1 = 0$.

Let $\mathsf{P}[X = -2] = \mathsf{P}[Y = +2] = 1/3$, $\mathsf{P}[X = +1] = \mathsf{P}[Y = -1] = 2/3$; then each has $\mu_1 = 0$ and $\mu_2 = 2$, but the distributions differ.