Problems for 2018 University of Texas Putnam Prep Session, week 6 (Oct 25)

Probability

1. Two coins are found in a fountain. One is a fair coin and the other has "heads" on both sides. One coin is chosen randomly and flipped 10 times. All 10 times it lands "heads" face up. What is the probability that the fair coin was chosen?

ANSWER: Let me answer this one slowly to help clarify the kind of reasoning that I find helpful when analyzing discrete probability questions.

The sides of the unfair coin may both look like heads but they are physically different surfaces. Make them look different by putting a dot on the head's eyes, say, so we can distinguish them as sides A and B. Then the result of choosing a coin and flipping it will be one of the outcomes in the following sample space (where F means "fair coin is chosen" and U means "unfair coin is chosen"):

$$\begin{aligned} S = \{&FHHHHHHHHHH, FHHHHHHHHHT, \ldots, FTTTTTTTTTT, \\ &UAAAAAAAAAA, \ldots, UBBBBBBBBBB\} \end{aligned}$$

This is a set of 2^{11} possible outcomes, all equally likely. (You may want to convince yourself of that last statement by viewing the elements of S as the leaves on a tree that branches in half 11 times as a result of asking 11 questions: first "which coin was chosen?"; then "what was the result of the first flip?"; and so on for each of the flips. As we move from the root to the leaves of the tree, at each stage we have a conditional probability of exactly 1/2 of getting either of the two answers to the question. Multiplying those conditional probabilities together as we move towards each leaf shows that the probability of ending on that leaf is exactly $1/2^{11}$.)

So the probability of any event $E \subseteq S$ is then $|E|/|S|$ (where I use absolute-value bars to indicate cardinality). Some events of interest here are

$$\begin{aligned} A &= \text{the unfair coin was chosen}; Pr(A) = 1024/2048 = 1/2 \\ A^c &= \text{the fair coin was chosen}; Pr(A^c) = 1 - Pr(A) = 1/2 \\ B &= \text{every flip showed a head}; Pr(B) = 1025/2048 \end{aligned}$$

Observe that in this problem $A \subseteq B$; in fact there is only one other element of S that lies in B besides the elements of A. That is, $Pr(A^c \cap B) = 1/2048$.

Now we can discuss the conditional probability $Pr(A^c|B)$ that was asked for. As implied in my discussion of trees, above, the general definition of conditional probabilities comes from the equation $Pr(E \cap F) = Pr(F) \cdot Pr(E|F)$, which allows us to compute them: $Pr(E|F) = Pr(E \cap F)/Pr(F)$. In our case this gives

$$Pr(A^c|B) = Pr(A^c \cap B)/Pr(B) = (1/2048)/(1025/2048) = 1/1025.$$

2. Sammy and Bevo each choose a real number at random between 1 and 10, inclusive. What is the probability that they differ by more than 4?

ANSWER: Our sample space is the square $S = \{(x,y) \,|\, x \in [1,10]$ and $y \in [1,10]$ (of area 81). Implied but not stated is the assumption that both players choose their number *with a uniform probability distribution*, meaning that the probability that the number they choose lies in a given interval is proportional to the *length* of that interval. It follows that the probability of any event $E \subseteq S$ is proportional to the area of E.

The event of interest to us is $E = \{(x,y) \,|\, |x-y| > 4\}$. This set E can be sketched within S: it's the set of points that are either above the line $y = x+4$ or below the line $y = x-4$, and within the square S these regions make two isosceles right triangles with legs of length 5. The sum of their areas is then 5^2, so the probability is

$$Pr(E) = \text{area}(E)/\text{area}(S) = 25/81.$$

3. There are 1,000 points equally spaced on a circle of radius 10. Six points are chosen randomly; call them A, B, C, D, E and F (in some order). What is the probability that the triangles ADC and BEF do not intersect each other?

ANSWER: Our sample space S is the set of sequences of six points chosen among the 1,000. (Tell me first which of the 1000 points will be called A, then tell me which of the remaining 999 points will be called B, etc.) So $|S| = 1000 \cdot 999 \ldots 995$.

But the same sample space can be thought of as a Cartesian product, that is, we write each element of S as an ordered pair (s, σ) where s is the *set* of points chosen and σ is a permutation of them — we scramble the six points (which can be listed in order of the point number, from 1 to 1000) into a new order to decide which is point A, which is B, etc. There are $\binom{1000}{6}$ such subsets and 6! such permutations, and the product of these agrees with the $|S|$ I computed above.

The reason to do this is that the question of whether the triangles intersect or not is independent of s and depends only on σ: replacing any point s with either of its neighbors on the circle will not make non-intersecting triangles intersect or vice versa, unless that neighboring point was another of the six poins in s.

Thus when we count the elements in S that do or do not correspond to intersecting triangles, we see we will count by multiples of that binomial coefficient; the probability $|E|/|S|$ will be $n/6!$ where n is the number of permutations σ that correspond to non-intersecting triangles.

In a similar way we can view the set of all these permutations as itself being a Cartesian product: simply tell me how to split the points into two subsets of three ($\{A, D, C\}$ versus $\{B, E, F\}$) and *then* tell me a permutation consistent with that split. By a "split" I mean to decide which of the six points will be in the same triangle as the lowest-numbered point (recall the points are numbered 1 to 1000); we must select 2 of the other five points and there are $\binom{5}{2} = 10$ ways to do so; whether we decide that first triangle will be called ADC or BEF will not affect whether the triangles cross or not, nor will our choice of which of the three points is called A or D or C, nor likewise the permutation of the other three vertices. So the $6! = 720$ permutations are now collected into 10 clusters of size

$2 \times 3! \times 3! = 72$ and each cluster behaves uniformly when we ask about intersections; if k of the clusters give non-intersecting triangles, the probability we seek is now $n/6! = k/10$.

So, finally, we examine these ten configuration, let's say with six points that are approximately near the vertices of a regular hexagon. It's easy to find 3 splits that keep the triangles apart: draw a line through any pair of opposite sides of the hexagon and use that line to split the six vertices into two groups. If the vertices are numbered in order around the hexagon those splits are $123|456, 156|234$, and $126|345$. A little picture will show that the other 7 configurations give intersecting triangles: $135|246$ is the "Star of David" configuration and $124|356, 125|236, 134|256, 136|245, 145|236$, and $146|235$ are all rotations and reflections of the same figure.

Thus $k = 3$ and the probability we seek is $3/10$.

4. All thirteen spades in a deck of cards are shuffled uniformly and dealt in a line. Let S be a statement about the order of the thirteen cards and $P(S)$ be the probability that S is true. For example, suppose S is "The five appears before the nine"; then $P(S) = 1/2$. How many of the values $1, 1/2, 1/3, \ldots, 1/50$ can $P(S)$ attain?

ANSWER: Don't get too hung up on "what is a statement?"; it is simply a way of attaching the words "true" or "false" to each ordering of the numbers from 1 to 13. The statement in the problem is one such function; another one would be "the order of the cards is one of the following: (blah,blah,. . .,blah), or (blah,blah,. . .,blah), or . . .". Irrespective of how you *word* such a statement, when you compute the probability, you simply need to decide which of the 13! orderings give a "true" to the statement S; that set of orderings is your event E. Then the probability that S will be a true statement whenever you run the experiment will, as always, be $|E|/13!$.

So the possible values of such a probability are all possible fractions in $[0, 1]$ which have a denominator dividing 13! (when expressed in lowest terms). In particular, the probability can be $1/n$ if and only if $n|13!$.

So this is really a number-theory problem! The question asks which integers up to 50 divide 13!. We can do this by direct computation, but more generally we can describe *all* the integers n that divide 13!. Do this using the prime factorizations of n and of 13!, say

$$n = \prod p^{e_p} \qquad \text{and} \qquad \prod p^{f_p}$$

where the products run over all primes (so that most of the exponents e_p and f_p are zero). Then $n|13!$ iff $e_p \leq f_p$ for all p.

For this purpose it's helpful to note that for any integer N,

$$N! = \prod p^{\lfloor n/p \rfloor + \lfloor n/p^2 \rfloor + \lfloor n/p^3 \rfloor + \ldots}$$

so in particular $13! = 2^{10} 3^5 5^2 7^1 11^1 13^1$ and thus the divisors of 13! are all the integers whose prime factorizations involve no higher exponents than these seven.

This includes all integers through 50 except these thirteen:

$$17, 19, 23, 29, 31, 34, 37, 38, 41, 43, 46, 47, 49$$

5. Shanille OKeal shoots free throws on a basketball court. She hits the first and misses the second, and thereafter the probability that she hits the next shot is equal to the proportion of shots she has hit so far. What is the probability that she hits exactly 50 of her first 100 shots?

ANSWER: I would view the sample space S as being the collection of all strings of 100 Hs and Ms that start with HM. (So $|S| = 2^{98}$.) In this problem it is especially helpful to create a tree as before with S as the collection of leaves on a tree that consists of 98 layers of binary branching, for then we can calculate the probability of arriving at any individual leaf as the product of the conditional probabilities of an H or an M on the branches that lead up to that leaf. For example, the probablility of obtaining outcome $HMHHHH\ldots HH$ is, by the statement of the problem, $1 \cdot 1 \cdot (1/2) \cdot (2/3) \cdot (3/4) \ldots (98/99) = 1/99$; the probability of outcome $HMHMHM\ldots HM$ is

$$1 \cdot 1 \cdot (1/2) \cdot (1/3) \cdot (2/4) \cdot (2/5) \ldots (50/99) \approx (1/2)^{98}$$

We are concerned with event E, the set of outcomes (leaves) that include precisely 50 Hs (and 50 Ms). The last example I just gave is one of these. We must add the probabilities of each of these outcomes together to get the probability of the event E.

How many such outcomes are there? We need to place precisely 49 more Hs into the string, and can choose any of the 98 spots (after the initial "HM") to do so. There are precisely $\binom{98}{49}$ ways to do so.

Now, the different outcomes in S are *not* all equally likely (as my two examples above show), but just among the outcomes in E they are! Indeed, the pattern of the product of the conditional probabilities is clear: if the kth symbol in the outcome is (say) an M then the probability of getting that symbol is by the problem statement equal to $1/(k-1)$ times the number of occurrences of that symbol earlier in the string, e.g. if there are 7 misses prior to the 13th shot, then there must have been 5 hits, so the probability of a hit is $5/12$, and thus the probability of a miss is $7/12$. So in order to compute the probability of a particular outcome we multiply these fractions, whose denominators multiply to $99!$, and whose numerators include the products $1 \cdot 2 \cdot 3 \ldots$ of the numerators corresponding to the first, second, third,… occurences of an H, as well as the products $1 \cdot 2 \cdot 3 \ldots$ of the numerators corresponding to the first, second, third,… occurences of an M. Obviously the probability of any particular outcome is then $h!m!/99!$ where h is the number of hits among the last 98 shots, and similarly m is the number of misses.

In the outcomes of importance to us, then, the probability of every individual leaf is $49!^2/99!$. We then multiply by the number $\binom{98}{49}$ of such outcomes to get the probability of the event: $Pr(E) = 1/99$.

6. You have coins $C_1, C_2, \ldots, C_n$. For each k, coin C_k is biased so that, when tossed, it has probability $1/(2k+1)$ of falling heads. If the n coins are tossed, what is the probability that the number of heads is odd? Express the answer as a rational function of n.

ANSWER: I would say the sample space S is the set of all strings s consisting of n Hs and Ts.

Now, in this problem the 2^n outcomes in the sample space are not equally likely. You might find it preferable to construct a sample space in which they ARE equally likely; for example, replace the kth coin with a spinner that has $2k + 1$ equally likely targets $H, T_1, T_2, \ldots T_{2k}$, and then construct S to consist of strings of symbols of length n, each symbol coming from a suitable set of options.

Instead, I will stick with my original S and compute more than just the probability of this one "even/odd" event. Let's compute the probability that the number of heads is any other number i (an integer in [0,n]).

This "number of heads" is an illustration of the concept of a *random variable*: a real-valued function defined on a sample space. In our case, the random variable $X =$ "number of heads": for each $s \in S$, $X(s)$ is the number of Hs in the string s.

You should stop for a moment and get a mental picture of a generic random variable χ; it partitions the whole domain (the sample space) into subsets — there's the set of outcomes s where $\chi(s) = 0$, and the set of outcomes where $\chi = 5$, and so on. Let's call these sets E_0 and E_5 and so on. Every outcome lies in precisely one of these sets $E_c \subseteq S$. In our problem, the question is to compute the probability of the event $E_1 \cup E_3 \cup E_5 \cup \ldots$.

The trick here is to associate to any random variable χ the polynomial $P_\chi(t) = \sum c_i t^i$ where $c_i = Pr(E_i)$ is the probability that the value of the random variable is i. (For today I am assuming that the random variable only takes on non-negative integer values, so that this P really is a polynomial.) So for example when you roll two dice the sample space might the 36 ordered pairs of numbers between 1 and 6; χ might be the random variable that reports the sum of the numbers shown on the tops of the dice. Then you might already know that for this example,

$$P_\chi(t) = \frac{1}{36}t^2 + \frac{2}{36}t^3 + \ldots + \frac{6}{36}t^7 + \ldots + \frac{1}{36}t^{12}$$

What is not immediately obvious until you think it through is that the polynomial I have just written out is the square of the polynomial $(1/6) \cdot (t + t^2 + t^3 + t^4 + t^5 + t^6)$. That's the polynomial that corresponds, in the same general way, to the random variable that reports the number shown on the top of ONE die. And this is a general pattern: if χ_1 and χ_2 are two independent random variables on a sample space, then you can add them (as you can add any functions on any domain!) to get another random variable $\chi = \chi_1 + \chi_2$ and — here's the punchline— $P_\chi(t) = P_{\chi_1}(t) \cdot P_{\chi_2}(t)$. Think that through: ask yourself which elements in S will give $\chi(s)$ the value of, say, 5; then ask what is the probability that an outcome s is in that set. It will be precisely the coefficient of t^5 when you expand this product of two polynomials.

So now let us return to our "number of heads" random variable X. This X is the sum of n very simple random variables X_k which just count the number of heads that showed up on the kth toss — which can obviously be only a 0 or a 1! (It's a useful trick: instead of asking *whether* something happened, ask *how many times* it happened, knowing that the answer is 0 or 1.) That's very handy: according to the preceding paragraph, that means $P_X(t) = P_{X_1}(t) \cdot P_{X_2}(t) \ldots P_{X_n}(t)$. And those factors are very easy to compute: since each X_k only takes on the values of 0 and 1, our definition of the Ps shows $P_{X_k}(t) = (1/(2k+1))t^1 + (2k/(2k+1))t^0$.

So here is the polynomial that corresponds to the random variable X:

$$P_X(t) = (\frac{1}{3}t + \frac{2}{3}) \cdot (\frac{1}{5}t + \frac{4}{5}) \dots (\frac{1}{2n+1}t + \frac{2n}{2n+1})$$

Simply expand out this polynomial, and read off the coefficients of the different powers of t; the coefficient of t^i is the probability that exactly i heads turn up when we perform all our coin tosses.

Fortunately we don't need any particular one of these coefficients. What we do want is the sum of all the odd coefficients. That's not hard to get. If P is any polynomial, $P(1)$ is the sum of its coefficients. Also $P(-1)$ is the alternating sum of the coefficients, starting with the coefficient of t^0. You can make one half or the other of the coefficients drop out by adding or subtracting these two quantities. Thus the probability we seek is $Pr(X = odd) = \frac{1}{2}(P_X(1) - P_X(-1))$

Now, for any random variable $P_\chi(1)$ is the sum of the coefficients, each of which is the probability that χ takes on one value or another; the sum of these is 1 since surely χ will take on *some* value! The other term $P_X(-1)$ is a bit harder but we may use the product formula above: $P_X(-1) = ((-1+2)/3) \cdot ((-1+4)/5) \dots ((-1+2n)/(2n+1))$. Both numerator and denominator are products of consecutive odd integers, so there is massive cancellation and $P_X(-1) = 1/(2n+1)$. Hence

$$Pr(X = odd) = \frac{1}{2}(P_X(1) - P_X(-1)) = \frac{1}{2}(1 - \frac{1}{2n+1}) = \frac{n}{2n+1}$$

7. If α is an irrational number, $0 < \alpha < 1$, is there a finite game with an honest coin such that the probability of one player winning the game is α? (An honest coin is one for which the probability of heads and the probability of tails are both $1/2$. A game is finite if with probability 1 it must end in a finite number of moves.)

ANSWER: We can do this for any α.

First let me describe a "universal game". This one is NOT a finite game; indeed it never ends and has no winners. (It's a "cooperative game".) Two players take turns flipping a coin. After the nth coin toss they create an interval $I_n = [a_n, b_n]$, starting with $I_0 = [0, 1]$. If the nth coin toss comes up heads, then I_n will be the right half of I_{n-1}, otherwise it will be the left half. So these are nested intervals of geometrically decreasing lengths, and they intersect in a single real number $L = \sup\{a_n\} = \inf\{b_n\}$. Assuming the coin is fair, the values of L are uniformly distributed in $[0, 1]$.

So now, given a number α as above we will play the following game: begin to play the universal game. If at some point the number α is not in the interval I_n then the game is over; if α is to the right of I_n then player 1 is the winner, otherwise player 2 wins. This game CAN continue forever without stopping, but that would require $L = \alpha$, an event that occurs with probability zero. More precisely, after n coin tosses the players will have created an interval $[k/2^n, (k+1)/2^n]$ for some k, and only with probability $1/2^n$ will this interval contain α; in the remaining cases the game will have ended with an interval not containing α. There are $\lfloor 2^n\alpha \rfloor$ intervals to the left of α, each of which will have been created with probability $1/2^n$, so the probability of player 1 winning by the nth round is $\lfloor 2^n\alpha \rfloor/2^n$. Apart from the interval containing α, that leaves the other intervals as a win for player 2, with probability $(2^n - 1 - \lfloor 2^n\alpha \rfloor)/2^n$. As n increases, the probability of the game continuing drops to zero and the probability of a win for player 1 approaches α, as desired.