STATISTICS 244 **Problem set 1** **9:30-10:50 TTh**

Due Thursday January 17, 2019

1. **The French Lottery.** Between April 1758 and May 1836, the French conducted a lottery very much like modern Lotto. As often as three times a month there would be a Tirage (a draw) that consisted of a sequence of 5 integers drawn without replacement from 1, 2, 3, ..., 90. The result of a "typical" Tirage might look like this: 61, 4, 25, 1, 71. The rules for the French lottery differed slightly from modern Lotto: before the Tirage a player would select one, two, three, four, or five numbers between 1 and 90, and had a choice of bets. If the player selected a single number (called an "Extrait"), the player is betting that number will be among the five in the Tirage. If two numbers are specified (an "Ambe") the bet is that the pair will be included in the Tirage. Similarly if three are specified (a "Terne"), four (a "Quaterne"), or five (a "Quine"). There was one further wrinkle in the betting: if the player bets an Extrait, there was a choice of two bets: a "Extrait simple" (the player would win if the number occurred in any of the five positions), or an "Extrait déterminé", where the player would also specify in which of the five positions the number would occur. Similarly, an "Ambe déterminé" would have the player specify for each member of the pair exactly where in the five draws of the Tirage it would occur. For example, the following bets would have won with the above "typical" Tirage: (i) if the number 71 were bet as an Extrait simple, (ii) if the numbers 1, 25, 71 were bet as a Terne, (iii) if the numbers 1, 4, 25, 61, 71 were bet as a Quine, (iv) if the player bet the Ambe déterminé that the second number selected would be 4 and the third would be 25. Clearly the order in which the numbers are drawn only matters with "déterminé" bets.

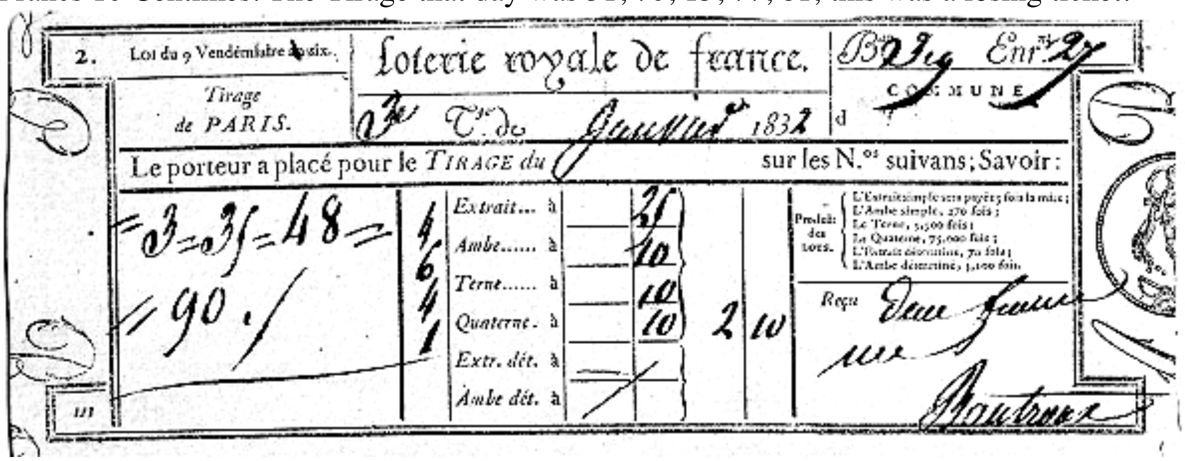
The payoffs would be a multiple of the bet placed, according to the following schedule. For example, if you bet \$5 on the Extrait simple "54", you would pay the lottery \$5 at the time of the bet. You would lose the \$5 if none of the five numbers in the Tirage were "54", but if a "54" came up in any of the five positions you would receive \$5x15=\$75 (but they would keep the \$5 you paid, so your net gain is \$70).

Type of Bet	Payoff: Multiple of Amount Bet
Extrait simple	15
Extrait déterminé	70
Ambe simple	270
Ambe déterminé	5,100
Terne	5,500
Quaterne	75,000
Quine	1,000,000

Assuming the Tirage is fairly conducted,
(a). How many different ways can a Tirage be selected (without regard to order)? Compute the answer (i) exactly, with a hand calculator, (ii) approximately, using Stirling's Formula (see Chapter 1 of Notes).
(b) What is the expected value (following Huygens) of a single bet of \$5 on an Ambe? A Terne? A Quaterne?
(c) The above Payoff values are not favorable to the player. Calculate what the payoff multiples should be for the Extrait Simple, Terne, and Quaterne if Huygens were to consider the bets fair. [Hint: The probability of winning a bet on an Extrait déterminé is clearly 1/90, since there are 90 possible outcomes, only one of which is favorable. The payoff should then be 90 times the amount bet in order that the bet be fair. To see this, suppose the bet is \$1 and the payoff is \$X. Then since with probability 1/90 you win X and with probability 89/90 you win \$0, the expected value is X(1/90) + (0)(89/90), which will equal the amount paid (\$1) only if X = 90. The other payoffs would be calculated similarly.]
(d) Suppose you placed multiple bets on the lottery, as you were permitted to do. In particular, suppose you specified these four numbers 16, 33, 35, 88 and bet on all possible Extraits, all Ambes, all Ternes, and the Quaterne involving these numbers (e.g. there are four possible Extraits). (i) What is the probability you will win at least one of these bets?

(ii) What is the probability that you will win all of these bets? (iii) Given that you win at least one Terne, what is the probability that you win the Quaterne?
(e) At various times the bet "Quine" was prohibited, principally to limit the potential loss due to a fraudulent ticket. However, an astute gambler could do even better (higher payoff and greater chance of winning) by making bets on successive lotteries. How might this be done? One method is to bet (say) \$1 on an Extrait simple, and if you win, in the next lottery bet all your winnings on another Extrait simple. Continue until you either lose or have won six times. Show that this works. Perhaps you can show this is the best such method – or not.

Figure: A betting slip from 1832 (when Quine's were not allowed), betting on the numbers 3, 35, 48, 90, with 25 Centimes bet on each of the four Extraits, 10 bet on each of the 6 Ambes, 10 bet on each of the four Ternes, and 10 bet on the Quaterne, for a total of 2 Francs 10 Centimes. The Tirage that day was 51, 70, 15, 77, 81; this was a losing ticket.



2. Huygens Proposition #3 can be extended to gambles with any finite number n of prizes. Here is a specific example for n=3. Suppose the possible prizes are a_1=\$1, a_2=\$2, a_3=\$15, and the numbers of chances for these prizes are p_1=10, p_2=4, p_3=2. Then,
(a) What is the fair price x for the gamble?
(b) Huygens proof can be adapted to this case. Suppose you are one of $16 = p_1 + p_2 + p_3$ people who each put up a stake of size x. Describe (exactly, in numerical terms) the side bets that will make this lottery equivalent to the gamble.

3. Use the definition of conditional probability to prove that for any two events defined on the same sample space, either $P(E|F) \le P(E) \le P(E|F^c)$ or $P(E|F^c) \le P(E) \le P(E|F)$.

4. Prove or Disprove: If $P(A) < P(B)$ and $P(C) > 0$, then $P(A|C) \le P(B|C)$.

5. In each of the following cases, state if the experiment can be considered as a Binomial experiment, and if so, (i) what are A and A^c, (ii) what are n and θ, and (iii) what is $P(X = 2)$.
(a) A coin that is weighted (with H twice as likely as T) is spun 6 times and the number X of H's is counted.
(b) A properly balanced European roulette wheel has 37 slots, one of them is green, 18 are red, 18 are black. The wheel is spun 10 times and X is the number of times a slot occurs that is not red.
(c) A survey researcher is taking a poll in Peoria to measure support for a proposed new airport and visits 50 households chosen at random from Census records. In each house if someone answers the door, the researcher speaks to each adult in turn, talking to a total of 45 adults among those where someone answers the door, and asks for their opinion (Supports, Opposes, or DK = don't know). Suppose that in fact 20% of the adults in Peoria are "DK" and everybody who answers, answers truthfully. Let X = # responses in the survey saying "DK" out of the 45 answers.

6. Monte Hall Game. Consider the Monte Hall Game, as described on the second handout. There it is convincingly argued that when presented with a choice, you should switch doors. But from one point of view, insufficient information is given for that conclusion. In that book's analysis it is implicitly assumed that the host will always offer a choice. Let us call a host who always offers a choice a "Type A" host. But what if the host were of "Type B", namely he only offers a choice if he knows you have selected the right door? With a Type B host you would, with probability 2/3, not be faced with any decision at all. But if you were faced with a choice, you surely should not switch! Now let us complicate things. Suppose the game show chooses a host at random, picking Type A with probability p and Type B with probability $1 - p$. Clearly if $p = 1$ you should switch and if $p = 0$ you should not switch. How large must p be in order that, when you are offered a choice, you should switch?

7. In 1693 Samuel Pepys wrote to Isaac Newton requesting advice related to a bet he had made. The question he posed involved three bets and asked which of the three bets has the greatest chance of winning. The bets involved the tossing of sets of fair dice (or "dyes" as Newton wrote in the following restatement). Newton's restatement:

> "What is the expectation or hope of A to throw every time one six at least with six dyes?
>
> "What is the expectation or hope of B to throw every time two sixes at least with twelve dyes?
>
> "What is the expectation or hope of C to throw every time three sixes at least with 18 dyes?"

Interpreting "hope or expectation" as "probability" (so bet A is "throw at least one '6' with a toss of six fair dice"), find the three probabilities of A, B, C exactly, both to two decimal places and as a ratio of integers. Newton did this correctly without a hand calculator; you may use any such aids.

[Bonus Material!]
Why it is hard to precisely evaluate the Expected Value of the Illinois Lottery

Illinois Lotto. [`www.illinoislottery.com`] Illinois Lotto is simpler than the earlier French version in the way it is conducted, but the payoffs are more complicated. In the Illinois version, 6 numbers are drawn without duplication from the integers from 1 to 52 (so it is called a 52/6 lottery). So a draw (English for "Tirage") might look like this: 34, 12, 17, 46, 47, 5. A single play consists of the bettor also selecting 6 numbers, and the player is a Grand Prize winner if the numbers bet are the same as those drawn (without regard to the order in which they are drawn). The price for one play is now \$1.00. If the player matches only 5 of the 6 numbers drawn the player is awarded a 2nd prize; a match of 4 is good for a 3rd prize and a match of 3 for a 4th prize. Otherwise the player loses. The prizes are all based upon the amount bet on that draw across the state, with the Grand Prize increased by whatever was left from the Grand Prize pool from the previous game of Lotto (if no one won a Grand Prize that time). Here are the rules that make this a complicated game to evaluate. These rules below are from 2017, when a play cost only \$0.50; there have been a few other changes.

(a) Half (50%) of the money taken in is assigned to the Winning Pool, money to be awarded as prizes, with that 50% to be allocated (see below) among the 4 prize categories. The other half taken in goes to the State of Illinois as profit, which is assigned to overhead and state education.
(b) Of the amount in the Winning Pool, 69.5% is assigned to the Grand Prize category (where it is combined with any money carried over from the previous drawing if no one won then).
(c) 5.5% of the Winning Pool is reserved to be split among 2nd prize winners.
(d) 10% is reserved for 3rd prize winners.
(e) 15% is for 4th prize winners (however they are guaranteed at least \$3.00 each, with the state coffers presumably making up any discrepancy).
(f) Grand Prize winners split the available pool, so that if there is (say) \$12,000,000 in the pool and three people matched the 6 numbers, each winner would get \$4,000,000. However a Grand Prize winner is paid that sum in 26 equal annual payments (of \$4,000,000/26 = \$153,846) over 26 years. If the winner wishes a lump sum payout at once, the amount paid is only the amount the state would have to pay a bank for an annuity to guarantee the 26 payments. Recently that was about 51% of the announced prize (\$4,000,000x.51 = \$2,040,000 in the example). (This corresponds to an annual interest rate of about 6.5%.)

Suppose that 10,000,000 "plays" are sold for a game of Lotto (at the 2017 price of \$0.50 each), and that \$4,000,000 is carried over from the previous game (since that game's prize was unclaimed).
What is the expected value of a single play in 2017?

First, let us compute the sizes of the various Prize Pools. Since we are told that 10,000,000 "plays" were sold at \$0.50 each, a total of \$5,000,000 was collected. Of this half or \$2,500,000 goes to the combined Pool. Of this, 69.5% or \$1,737,500 goes to the Grand Prize Pool, where it is combined with the \$4,000,000 that was carried over, making a Grand Prize Pool (for those who match all 6) of $M(6) = \$5,737,500$. The 2nd Prize Pool (matching 5 only) is $M(5) = \$2,500,000 \times 0.055 = \$137,500$. The 3rd Prize Pool (matching 4) is $M(4) = \$2,500,000 \times 0.10 = \$250,000$. And the 4th Prize Pool (matching 3) is $M(3) = \$2,500,000 \times 0.15 = \$375,000$.

Next consider the chances that you will win the various prizes. You will only win one prize, the highest level one that you are eligible for (so if you match all 6 numbers you will win only the Grand Prize).

There are $\binom{52}{6}$ = 20,358,520 possible different draws. Of these, the number of ways you can choose exactly k that match the winning number is $\binom{6}{k}\binom{46}{6-k}$. The chance of matching exactly k is therefore $P(k) = \binom{6}{k}\binom{46}{6-k}\Big/\binom{52}{6}$. Using a computer or Stirling's Formula you can find that P(6) =4.912 x 10^{-8} = 0.00000004912, P(5) = 1.3557 x 10^{-5} = 0.000013557, P(4) = 0.0007626, and P(3) = 0.0149. There is a temptation to give the expected winnings as M(6)xP(6) + M(5)xP(5) + M(4)xP(4) + M(3)xP(3), but that would only be correct if you could count the Grand Prize at full value and if you did not have to split any winning pool, and neither of these is correct. Instead of $5,737,500, the present value of that prize (using the figure provided) is only 51% or $M^*(6)$= $2,983,500. But how to think about how many ways the pools are split? The number of people you would share with if you win is a random variable. Let N(k) be the number of <u>other</u> winners you would have to share the k^{th} prize with if you are lucky enough to win the k^{th} prize, that is the number of <u>other</u> bets that also win that prize. Then apparently your expected winnings would be something like M(6)xP(6)/(N(6)+1) + M(5)xP(5)/(N(5)+1) + M(4)xP(4)/(N(4)+1) + M(3)xP(3)/(N(3)+1). But this is random! To evaluate its expectation we must make some assumptions about other people's betting behavior. Let us assume they all guess randomly (this is called using a "Quickpick"), independently of one another. There are 9,999,999 "other" bets and then N(k) will have a Binomial (n=9,999,999, θ=P(k)) distribution. To find your true expected winnings it in then necessary to evaluate the expectation of 1/(N(k)+1) for k = 6, 5, 4, 3. This is beyond the scope of our present tools, but a pretty good approximation is 1/(9,999,999xP(k)+1). Using this we find the expectation of a single lottery pick to be $M^*(6)$xP(6)/(9,999,999xP(6)+1) + M(5)xP(5)/(9,999,999xP(5)+1) + M(4)xP(4)/9,999,999xP(4)+1) + M(3)xP(3)/(9,999,999xP(3)+1) = .098+.014+.025+.037 = $0.174. Since a bet costs $0.50, this suggests the lottery is not a very good deal. This calculation depends upon a number of assumptions (like the number of bets made and the size of the carryover) that would vary from day to day and would be hard to estimate in advance.