CS 70 Spring 2017 — Discrete Mathematics and Probability Theory — Rao

HW 12

1 Sundry

Before you start your homework, write down your team. Who else did you work with on this homework? List names and email addresses. (In case of homework party, you can also just describe the group.) How did you work on this homework? Working in groups of 3-5 will earn credit for your "Sundry" grade.

Please copy the following statement and sign next to it:

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up.

I certify that all solutions are entirely in my words and that I have not looked at another student's solutions. I have credited all external sources in this write up. (Signature here)

2 Quadratic Regression

In this question, we will find the best quadratic estimator of Y given X. First, some notation: let μ_i be the ith moment of X, i.e. $\mu_i = \mathbf{E}[X^i]$. Also, define $\beta_1 = \mathbf{E}[XY]$ and $\beta_2 = \mathbf{E}[X^2Y]$. For simplicity, we will assume that $\mathbf{E}[X] = \mathbf{E}[Y] = 0$ and $\mathbf{E}[X^2] = \mathbf{E}[Y^2] = 1$. (Note that this poses no loss of generality, because we can always transform the random variables by subtracting their means and dividing by their standard deviations.) We claim that the best quadratic estimator of Y given X is

$$\hat{Y} = \frac{1}{\mu_3^2 - \mu_4 + 1}(aX^2 + bX + c)$$

where

$$\begin{aligned} a &= \mu_3\beta_1 - \beta_2, \\ b &= (1-\mu_4)\beta_1 + \mu_3\beta_2, \\ c &= -\mu_3\beta_1 + \beta_2. \end{aligned}$$

Your task is to prove the Projection Property for $\hat{Y}$.

(a) Prove that $\mathbf{E}[Y - \hat{Y}] = 0$.

(b) Prove that $\mathbf{E}[(Y - \hat{Y})X] = 0$.

(c) Prove that $\mathbf{E}[(Y - \hat{Y})X^2] = 0$.

Any quadratic function of X is a linear combination of 1, X, and X^2. Hence, these equations together imply that $Y-\hat{Y}$ is orthogonal to any quadratic function of X, and so $\hat{Y}$ is the best quadratic estimator of Y.

Solution:

(a) By linearity of expectation:

$$\mathbf{E}[Y-\hat{Y}] = \mathbf{E}[Y] - \frac{a\mathbf{E}[X^2]+b\mathbf{E}[X]+c}{\mu_3^2-\mu_4+1} = \frac{-a-c}{\mu_3^2-\mu_4+1} = 0$$

since $\mathbf{E}[X]=\mathbf{E}[Y]=0$ and $\mathbf{E}[X^2]=1$.

(b)

$$\begin{aligned}\mathbf{E}[(Y-\hat{Y})X] &= \mathbf{E}[XY] - \frac{a\mathbf{E}[X^3]+b\mathbf{E}[X^2]+c\mathbf{E}[X]}{\mu_3^2-\mu_4+1} \\ &= \beta_1 - \frac{(\mu_3\beta_1-\beta_2)\mu_3+((1-\mu_4)\beta_1+\mu_3\beta_2)}{\mu_3^2-\mu_4+1}\end{aligned}$$

which, after a little algebra, gives 0.

(c)

$$\begin{aligned}\mathbf{E}[(Y-\hat{Y})X^2] &= \mathbf{E}[X^2Y] - \frac{a\mathbf{E}[X^4]+b\mathbf{E}[X^3]+c\mathbf{E}[X^2]}{\mu_3^2-\mu_4+1} \\ &= \beta_2 - \frac{\mu_4(\mu_3\beta_1-\beta_2)+\mu_3((1-\mu_4)\beta_1+\mu_3\beta_2)-\mu_3\beta_1+\beta_2}{\mu_3^2-\mu_4+1}\end{aligned}$$

which, after a little algebra, gives 0.

3 Projection Property

Use the Projection Property to answer the following questions.

(a) Prove or disprove: for any function ϕ, $\mathbf{E}[\mathbf{E}[Y \mid X]\phi(X)] = 0$.

(b) Prove or disprove: $\mathbf{E}[(Y-\mathbf{E}[Y \mid X])L[Y \mid X]] = 0$.

(c) Prove the following: $\mathbf{E}[X^2 \mid Y] = \mathbf{E}[(X-\mathbf{E}[X \mid Y])^2 \mid Y] + \mathbf{E}[X \mid Y]^2$. [*Hint*: In the expression $\mathbf{E}[X^2 \mid Y]$, try replacing X with $(X-\mathbf{E}[X \mid Y])+\mathbf{E}[X \mid Y]$.]

(d) We have already shown that $\mathbf{E}[\mathbf{E}[Y \mid X]] = \mathbf{E}[Y]$. Prove that $\mathbf{E}[L[Y \mid X]] = \mathbf{E}[Y]$.

(e) Prove the following property of conditional expectation:

$$\mathbf{E}[\mathbf{E}[Z \mid X,Y] \mid X] = \mathbf{E}[Z \mid X].$$

[*Hint*: Take a closer look at the method by which we prove properties of conditional expectation in Note 26.]

Solution:

(a) This is *not* the Projection Property. As a reminder, the Projection Property states that

$$\mathbf{E}[(Y-\mathbf{E}[Y \mid X])\phi(X)] = 0.$$

For a simple counterexample, take $\phi(X) = \mathbf{E}[Y \mid X]$. Then, the statement reads

$$\mathbf{E}[\mathbf{E}[Y \mid X]^2] = 0,$$

which is only true if $\mathbf{E}[Y \mid X] = 0$.

(b) True, by the Projection Property: $L[Y \mid X]$ is a function of X, and $Y - \mathbf{E}[Y \mid X]$ is orthogonal to every function of X.

(c) Follow the hint:

$$\begin{aligned}
\mathbf{E}[X^2 \mid Y] &= \mathbf{E}[(X-\mathbf{E}[X \mid Y]+\mathbf{E}[X \mid Y])^2 \mid Y] \\
&= \mathbf{E}[(X-\mathbf{E}[X \mid Y])^2 \mid Y] + 2\mathbf{E}[(X-\mathbf{E}[X \mid Y])\mathbf{E}[X \mid Y] \mid Y] \\
&\qquad + \mathbf{E}[\mathbf{E}[X \mid Y]^2 \mid Y]
\end{aligned}$$

Examine the expression above. The second term is 0, because

$$\begin{aligned}
\mathbf{E}[(X-\mathbf{E}[X \mid Y])\mathbf{E}[X \mid Y] \mid Y] &= \mathbf{E}[X \mid Y]\mathbf{E}[X-\mathbf{E}[X \mid Y] \mid Y] = \mathbf{E}[X \mid Y](\mathbf{E}[X \mid Y]-\mathbf{E}[X \mid Y]) \\
&= 0.
\end{aligned}$$

For the last term, note that $\mathbf{E}[X \mid Y]$ is a function of Y, so $\mathbf{E}[X \mid Y]^2$ is also a function of Y. Conditioned on Y, any function of Y is effectively a "constant", so $\mathbf{E}[\mathbf{E}[X \mid Y]^2 \mid Y] = \mathbf{E}[X \mid Y]^2$. Hence,

$$\mathbf{E}[X^2 \mid Y] = \mathbf{E}[(X-\mathbf{E}[X \mid Y])^2 \mid Y] + \mathbf{E}[X \mid Y]^2.$$

(d) This follows from the Projection Property as well. $\mathbf{E}[Y - L[Y \mid X]] = 0$, so $\mathbf{E}[L[Y \mid X]] = \mathbf{E}[Y]$.

(e) We will use Lemma 26.1(b) from the course notes, which is reproduced below:

Lemma: If $g(X)$ has the property that

$$\mathbf{E}[(Y-g(X))\phi(X)] = 0 \qquad \forall \phi(\cdot),$$

then $g(X) = \mathbf{E}[Y \mid X]$.

The meaning of the lemma is that the Projection Property uniquely characterizes the conditional expectation. Hence, to show that $g(X) = \mathbf{E}[Z \mid X]$ is the conditional expectation of $\mathbf{E}[Z \mid X, Y]$ given X, it suffices to show that

$$\mathbf{E}[(\mathbf{E}[Z \mid X, Y] - g(X))\phi(X)] = 0 \qquad \forall \phi(\cdot).$$

Recall the law of iterated expectation, $\mathbf{E}[\mathbf{E}[Z \mid X]] = \mathbf{E}[Z]$.

$$\begin{aligned}
\mathbf{E}[(\mathbf{E}[Z \mid X,Y] - g(X))\phi(X)] &= \mathbf{E}[\mathbf{E}[Z\phi(X) \mid X,Y] - g(X)\phi(X)] && \phi(X) \text{ is a function of } X,Y \\
&= \mathbf{E}[Z\phi(X)] - \mathbf{E}[g(X)\phi(X)] && \text{iterated expectation} \\
&= \mathbf{E}[Z\phi(X)] - \mathbf{E}[\mathbf{E}[Z \mid X]\phi(X)] && \text{definition of } g(X) \\
&= \mathbf{E}[Z\phi(X)] - \mathbf{E}[\mathbf{E}[Z\phi(X) \mid X]] && \phi(X) \text{ is a function of } X \\
&= \mathbf{E}[Z\phi(X)] - \mathbf{E}[Z\phi(X)] && \text{iterated expectation} \\
&= 0.
\end{aligned}$$

So, $\mathbf{E}[\mathbf{E}[Z \mid X,Y] \mid X] = \mathbf{E}[Z \mid X]$. This is sometimes known as the tower property of conditional expectation.

4 Balls in Bins Estimation

We throw $n > 0$ balls into $m \geq 2$ bins. Let X and Y represent the number of balls that land in bin 1 and 2 respectively.

(a) Calculate $\mathbf{E}[Y \mid X]$. [*Hint*: Your intuition may be more useful than formal calculations.]

(b) What are $L[Y \mid X]$ and $Q[Y \mid X]$ (where $Q[Y \mid X]$ is the best quadratic estimator of Y given X)? [*Hint*: Your justification should be no more than two or three sentences, no calculations necessary! Think carefully about the meaning of the MMSE.]

(c) Unfortunately, your friend is not convinced by your answer to the previous part. Compute $\mathbf{E}[X]$ and $\mathbf{E}[Y]$.

(d) Compute $\text{var}(X)$.

(e) Compute $\text{cov}(X,Y)$.

(f) Compute $L[Y \mid X]$ using the formula. Ensure that your answer is the same as your answer to part (b).

Solution:

(a) $\mathbf{E}[Y \mid X = x] = (n-x)/(m-1)$, because once we condition on x balls landing in bin 1, the remaining $n-x$ balls are distributed uniformly among the other $m-1$ bins. Therefore,

$$\mathbf{E}[Y \mid X] = \frac{n-X}{m-1}.$$

(b) We showed that $\mathbf{E}[Y \mid X]$ is a linear function of X. Since $\mathbf{E}[Y \mid X]$ is the best *general* estimator of Y given X, it must also be the best *linear* and *quadratic* estimator of Y given X, i.e. $\mathbf{E}[Y \mid X]$, $L[Y \mid X]$, and $Q[Y \mid X]$ all coincide.

(c) Let X_i be the indicator that the ith ball falls in bin 1. Then, $X = \sum_{i=1}^n X_i$, and by linearity of expectation, $\mathbf{E}[X] = \sum_{i=1}^n \mathbf{E}[X_i] = n/m$, since there are n indicators and each ball has a probability $1/m$ of landing in bin 1. By symmetry, $\mathbf{E}[Y] = n/m$ as well.

(d) The number of balls that falls into the first bin is binomially distributed with parameters n and $1/m$. Hence the variance is $n(1/m)(1-1/m)$.

(e) Let X_i be as before, and let Y_i be the indicator that the ith ball falls into bin 2.

$$\operatorname{cov}(X,Y) = \sum_{i=1}^n \sum_{j=1}^n \operatorname{cov}(X_i, Y_j)$$

We can compute $\operatorname{cov}(X_i,Y_i) = \mathbf{E}[X_iY_i] - \mathbf{E}[X_i]\mathbf{E}[Y_i] = 0 - (1/m)(1/m) = -1/m^2$ (note that $\mathbf{E}[X_iY_i] = 0$ because it is impossible for a ball to land in both bins 1 and 2). Also, we have $\operatorname{cov}(X_i,Y_j) = 0$ because the indicator for the ith ball is independent of the indicator for the jth ball when $i \neq j$. Hence, $\operatorname{cov}(X,Y) = n(-1/m^2) = -n/m^2$.

(f)

$$\begin{aligned}
L[Y \mid X] &= \mathbf{E}[Y] + \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(X)}(X - \mathbf{E}[X]) \\
&= \frac{n}{m} + \frac{-n/m^2}{n(1/m)(1-1/m)}\left(X - \frac{n}{m}\right) \\
&= \frac{n}{m} - \frac{1}{m-1}\left(X - \frac{n}{m}\right) \\
&= \frac{mn - n - mX + n}{m(m-1)} = \frac{n-X}{m-1}
\end{aligned}$$

5 Swimsuit Season

In the swimsuit industry, it is well-known that there is a "swimsuit season". During this time, swimsuit sales skyrocket!

We will model this with a random variable X which is either λ_L or λ_H with equal probability; λ_L represents the mean number of customers in a day when swimsuits are not in season, and λ_H represents the mean number of customers during swimsuit season. So, λ_L is the "low rate" and λ_H is the "high rate". The number of customer arrivals Y on a particular day is modeled as a Poisson random variable with mean X.

You observe Y customers on a certain day, and the task is to estimate X.

(a) What is $L[X \mid Y]$?

(b) What is $\mathbf{E}[X \mid Y]$?

Solution:

(a) The key idea here is that X gives information about Y, so to calculate any quantity involving Y, it is helpful to condition on X.

First, we observe that since X is the mean of Y, $\mathbf{E}[Y \mid X] = X$. So,

$$\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[Y \mid X]] = \mathbf{E}[X] = \frac{1}{2}(\lambda_L + \lambda_H).$$

Now, we compute $\operatorname{cov}(X,Y)$.

$$\begin{aligned}\operatorname{cov}(X,Y) &= \mathbf{E}[XY] - \mathbf{E}[X]\mathbf{E}[Y] = \mathbf{E}[\mathbf{E}[XY \mid X]] - \mathbf{E}[X]^2 = \mathbf{E}[X\mathbf{E}[Y \mid X]] - \mathbf{E}[X]^2 \\ &= \mathbf{E}[X^2] - \mathbf{E}[X]^2 = \frac{1}{2}(\lambda_L^2 + \lambda_H^2) - \frac{1}{4}(\lambda_L + \lambda_H)^2 = \frac{1}{4}(\lambda_H^2 - 2\lambda_L\lambda_H + \lambda_L^2) \\ &= \frac{1}{4}(\lambda_H - \lambda_L)^2.\end{aligned}$$

Next is $\operatorname{var}(Y)$.

$$\operatorname{var}(Y) = \mathbf{E}[Y^2] - \mathbf{E}[Y]^2 = \mathbf{E}[\mathbf{E}[Y^2 \mid X]] - \mathbf{E}[\mathbf{E}[Y \mid X]]^2 = \mathbf{E}[X^2 + X] - \mathbf{E}[X]^2$$

(remember that if Y has the Poisson distribution with mean λ, then $\mathbf{E}[Y^2] = \lambda^2 + \lambda$)

$$= \frac{1}{2}(\lambda_L^2 + \lambda_H^2) + \frac{1}{2}(\lambda_L + \lambda_H) - \frac{1}{4}(\lambda_L + \lambda_H)^2 = \frac{1}{4}(\lambda_H - \lambda_L)^2 + \frac{1}{2}(\lambda_L + \lambda_H).$$

Hence,

$$\begin{aligned}L[X \mid Y] &= \mathbf{E}[X] + \frac{\operatorname{cov}(X,Y)}{\operatorname{var}(Y)}(Y - \mathbf{E}[Y]) \\ &= \frac{1}{2}(\lambda_L + \lambda_H) + \frac{(\lambda_H - \lambda_L)^2/4}{(\lambda_H - \lambda_L)^2/4 + (\lambda_L + \lambda_H)/2}\left(Y - \frac{1}{2}(\lambda_L + \lambda_H)\right).\end{aligned}$$

(b) We calculate $\Pr(X = \lambda_L \mid Y = y)$.

$$\begin{aligned}\Pr(X = \lambda_L \mid Y = y) &= \frac{\Pr(Y = y \mid X = \lambda_L)\Pr(X = \lambda_L)}{\Pr(Y = y \mid X = \lambda_L)\Pr(X = \lambda_L) + \Pr(Y = y \mid X = \lambda_H)\Pr(X = \lambda_H)} \\ &= \frac{e^{-\lambda_L}\lambda_L^y/y!}{e^{-\lambda_L}\lambda_L^y/y! + e^{-\lambda_H}\lambda_H^y/y!} = \frac{e^{-\lambda_L}\lambda_L^y}{e^{-\lambda_L}\lambda_L^y + e^{-\lambda_H}\lambda_H^y}.\end{aligned}$$

Therefore,

$$\mathbf{E}[X \mid Y = y] = \lambda_L \cdot \frac{e^{-\lambda_L}\lambda_L^y}{e^{-\lambda_L}\lambda_L^y + e^{-\lambda_H}\lambda_H^y} + \lambda_H \cdot \frac{e^{-\lambda_H}\lambda_H^y}{e^{-\lambda_L}\lambda_L^y + e^{-\lambda_H}\lambda_H^y}$$

so

$$\mathbf{E}[X \mid Y] = \frac{e^{-\lambda_L}\lambda_L^{Y+1} + e^{-\lambda_H}\lambda_H^{Y+1}}{e^{-\lambda_L}\lambda_L^Y + e^{-\lambda_H}\lambda_H^Y}.$$

This is a very non-linear function of Y, which illustrates that in general, the MMSE does not equal the LLSE.

6 Political War

Initially, there are d Democrats and r Republicans in a room. They begin to argue. On each day, a random person in the room leaves and returns with an additional member of his or her political party; that is, either a Democrat will leave and return with a Democrat friend, or a Republican will leave and return with a Republican friend. Let D_n denote the number of democrats in the room at the end of the nth day. Let $D_0 = d$.

(a) Find $\mathbf{E}[D_n \mid D_{n-1}]$.

(b) Find $\mathbf{E}[D_n]$ using the law of iterated expectation.

(c) What is the expected fraction of Democrats in the room at the end of day n?

Solution:

(a) Initially, there are $d+r$ people in the room, so at the start of day n, there are $n+d+r-1$ people in the room. If there are D_{n-1} Democrats at the start of day n, then the probability that a Democrat leaves the room is $D_{n-1}/(n+d+r-1)$; then, $D_n = D_{n-1}+1$. So, we can write $D_n = D_{n-1} + 1_{A_n}$, where 1_{A_n} is the indicator that a Democrat leaves the room on day n.

$$\mathbf{E}[D_n \mid D_{n-1}] = \mathbf{E}[D_{n-1} + 1_{A_n} \mid D_{n-1}] = D_{n-1} + \frac{D_{n-1}}{n+d+r-1} = \frac{n+d+r}{n+d+r-1} \cdot D_{n-1}.$$

(b) We have

$$\mathbf{E}[D_n] = \mathbf{E}[\mathbf{E}[D_n \mid D_{n-1}]] = \frac{n+d+r}{n+d+r-1} \cdot \mathbf{E}[D_{n-1}] = \frac{n+d+r}{n+d+r-1} \frac{n+d+r-1}{n+d+r-2} \cdot \mathbf{E}[D_{n-2}]$$
$$= \cdots = \frac{n+d+r}{n+d+r-1} \cdots \frac{d+r+1}{d+r} \cdot \mathbf{E}[D_0].$$

Observe that the product is telescoping: we can cancel out factors to obtain

$$\mathbf{E}[D_n] = \frac{n+d+r}{d+r} \cdot d$$

(remember that $D_0 = d$).

(c) At the end of day n, there are $n+d+r$ people in the room, so the expected fraction of Democrats is $d/(d+r)$. This is the exactly the same as the original fraction of Democrats in the room! This problem is more commonly known as Polya's urn problem.

Here is an interesting thought to ignite your curiosity: if we start with a million Democrats and one Republican in the room, then it is overwhelmingly likely for a Democrat to leave the room. The Democrat will return with yet another Democrat, which only makes it more unlikely that the number of Republicans in the room will increase. Over time, we might predict that the fraction of Democrats in the room will approach 1. Indeed, but in expectation this is simply not true.

7 Optimal Gambling

In even-money gambling games, you bet a fixed amount of money. If you win the game, you are given back the money that you bet, and you receive an additional amount of money equal to your original bet. If you lose the game, you lose the amount of money you bet.

(a) You are gambling and your probability of winning, on each round, is $1/2 < p < 1$: the game is in your favor! You use the following strategy: on each round, you will bet a fraction q of the money you have at the start of the round. Let X_n denote the amount of money you have on round n. X_0 represents your initial assets and is a constant value. What is $\mathbf{E}[X_n]$?

(b) What value of q will maximize $\mathbf{E}[X_n]$? For this value of q, what is the distribution of X_n? Can you predict what will happen as $n \to \infty$? [*Hint*: Under this betting strategy, what happens if you ever lose a round?]

(c) The problem with the previous approach is that we were too concerned about expected value, so our gambling strategy was too extreme. Let's start over: again we will use a gambling strategy in which we bet a fraction q of our money at each round. Express X_n in terms of n, q, X_0, and W_n, where W_n is the number of rounds you have won up until round n. [*Hint*: Does the order in which you win the games affect your profit?]

(d) By the law of large numbers, $W_n/n \to p$ as $n \to \infty$. Using this fact, what does $(\ln X_n)/n$ converge to as $n \to \infty$?

(e) The rationale behind $(\ln X_n)/n$ is that if $(\ln X_n)/n \to c$, where c is a constant, then that means for large n, X_n is roughly e^{cn}. Therefore, c is the asymptotic growth rate of your fortune! Find the value of q that maximizes your asymptotic growth rate.

(f) Using the value of q you found in the previous part, compute $\mathbf{E}[X_n]$.

Solution:

(a) At the start of round n, the amount of money you have is X_n. With probability $1-p$, you will lose the round and $X_{n+1} = (1-q)X_n$. With probability p, you will win the round and $X_{n+1} = (1-q)X_n + 2qX_n = (1+q)X_n$. Therefore,

$$\mathbf{E}[X_{n+1} \mid X_n] = (1-p)(1-q)X_n + p(1+q)X_n.$$

By the law of iterated expectation,

$$\mathbf{E}[X_{n+1}] = \mathbf{E}[\mathbf{E}[X_{n+1} \mid X_n]] = ((1-p)(1-q) + p(1+q))\mathbf{E}[X_n].$$

Therefore,

$$\mathbf{E}[X_n] = ((1-p)(1-q) + p(1+q))^n X_0.$$

(b) We want $(1-p)(1-q)+p(1+q)$ to be as large as possible. Note that this is linear in q, and the coefficient for q is $p-(1-p)>0$. Hence, we should take q to be as large as possible, which is 1 (you cannot bet more money than you actually have).

For this value of q, note that on each round you either double your money or go broke. Hence, the distribution is:

$$X_n = \begin{cases} 2^n X_0, & \text{with probability } p^n \\ 0, & \text{with probability } 1-p^n \end{cases}$$

Uh-oh. As $n \to \infty$, the probability that you are broke approaches 1. The issue here is that your expected fortune grows large, but the probability that you are rich grows vanishingly small. In general, $X_n \to 0$ as $n \to \infty$ does not necessarily imply that $\mathbf{E}[X_n] \to 0$, which is what we see here.

(c) You win W_n times and each time you win, your fortune is multiplied by $1+q$; you lose $n-W_n$ times, and each time you lose, your fortune is multiplied by $1-q$. Therefore,

$$X_n = X_0(1-q)^{n-W_n}(1+q)^{W_n}.$$

(d) One has

$$\begin{aligned} \frac{\ln X_n}{n} &= \frac{\ln X_0}{n} + \left(1 - \frac{W_n}{n}\right)\ln(1-q) + \frac{W_n}{n}\ln(1+q) \\ &\xrightarrow[n\to\infty]{} (1-p)\ln(1-q) + p\ln(1+q). \end{aligned}$$

(e) We can use calculus to optimize c:

$$\frac{\mathrm{d}}{\mathrm{d}q}((1-p)\ln(1-q)+p\ln(1+q)) = -\frac{1-p}{1-q} + \frac{p}{1+q}.$$

Set the derivative to 0:

$$\frac{p}{1+q} = \frac{1-p}{1-q} \implies p - pq = 1 + q - p - pq \implies q = 2p-1.$$

This is known as the Kelly betting criterion. The CS 70 course staff is not responsible for any losses you incur with this betting strategy, but we do think it's pretty cool that you can analyze optimal gambling with the methods in this course.

(f) We can plug in $q = 2p-1$ from our previous result:

$$\mathbf{E}[X_n] = ((1-p)(1-q)+p(1+q))^n X_0 = 2^n(p^2+(1-p)^2)^n X_0.$$