ECE 792-41 Statistical Foundations
for Signal Processing and Machine Learning (Fall 2020)
Instructor: Dr. Chau-Wai Wong

Homework 0
Topic: Probability Theory Review

Animal Crossing: New Horizons [1] is a video game that allows players to experience various aspects of real life in a simulated world. One interesting feature is its stalk market [2] mimicking real-world stock markets. The game allows players to buy turnips each Sunday morning and to sell them from Monday morning to Saturday night, for a total of 12 half-day trading sessions. According to a reverse engineering effort by Ninji, the selling price for each session is determined at the beginning of each session, and the change of the price follows a certain random process that was detailed in his/her GitHub code [3]. Players have to sell the turnips by the end of the last trading session, or they will go bad and cause a total loss in their turnip investment.

For each week, the pricing strategy of the 12 trading session follows one of four predefined random processes [3]. In this homework, we will investigate some interesting pricing strategy using basic probability tools that you have learned in your undergraduate probability/statistics courses. Prior knowledge about Animal Crossing is not needed. To simply the presentation, the price we refer to in this document is a price normalized by the base price.

Problem 1 (Second Best Session Price Distribution) Let us investigate the pricing strategy of the third session (the second best session) of the *increasing phase* of the *small spike pattern.* We adopt the terminologies from [2]. First, an intermediate price Y_0 is picked uniformly random from 1.4 to 2.0. Second, the session price X_0 is picked uniformly random from 1.4 to Y_0. Let us define $X = X_0 - 1.4$.

a) Show that $X \sim \text{Unif}(0, Y)$ and $Y \sim \text{Unif}(0, 0.6)$. We call them the offset-removed prices, and we will use them in the remainder of this document.

b) Show that the CDF, PDF, expectation of X are

$$F(x) = \frac{x}{0.6}\left[1 - \ln\left(\frac{x}{0.6}\right)\right], \text{ for } 0 < x < 0.6, \tag{1}$$

$$f(x) = -\frac{1}{0.6}\ln\left(\frac{x}{0.6}\right), \text{ for } 0 < x < 0.6, \tag{2}$$

$$E[X] = 0.15. \tag{3}$$

[Hint: You may either start from the definition of CDF, i.e., $F(x) = P[X \leq x]$, in which you may need $X = ZY$, $Z \sim \text{Unif}(0, 1)$; or start from the joint PDF, i.e., $f_{XY}(x, y)$.]

c) Figure 1 shows the PDF for $X_0 = X + 1.4$. Give intuitive explanation from the generation process of X_0 why it cannot be uniformly distributed.

Problem 2 (Best Session Price Posterior Distribution) Let us investigate the pricing strategy of the fourth session (the best session) of the *increasing phase* of the *small spike pattern.* We adopt the terminologies from [2]. The price at the fourth session equals to the intermediate price Y generated in the third session. Suppose a player has observed the price in the third session, namely, $X = x$.

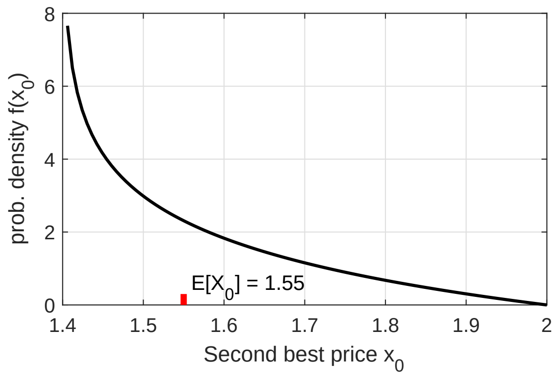

Figure 1: PDF $f_{X_0}(x_0)$ for the second best session price X_0. Note that this distribution is highly nonuniform. The mean value of the session price is 1.55 as indicated by the red bar on the x-axis.

a) What will the price distribution look like for the fourth session given the observed price in the third session? Show that the conditional distribution is

$$f_{Y|X}(y|x) = \frac{1}{y\,(\ln 0.6 - \ln x)}, \quad 0 < x < y < 0.6. \tag{4}$$

b) Show that the average price of the fourth day given the price from the third day takes the following form:

$$E[Y|X = x] = \frac{0.6 - x}{\ln 0.6 - \ln x}, \quad 0 < x < 0.6. \tag{5}$$

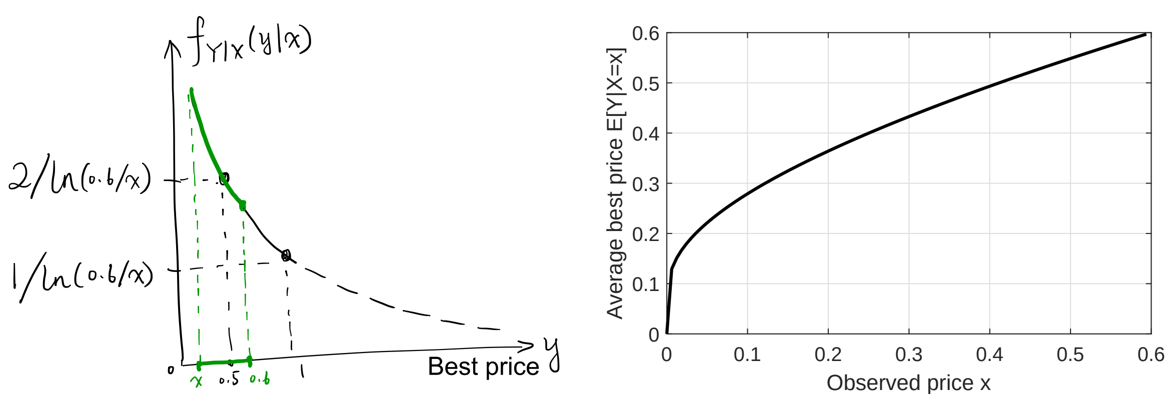

Figure 2: (Left) Conditional PDF $f_{Y|X}(y|x)$ for the best price of the fourth session Y given the observed price of the third section. It is more likely to get a lower price no matter what the observed price is. (Right) The conditional expectation for the best price of the fourth session given the observed price. When the observed price is greater than 0.1, the expected price is almost increasing linearly in the observed price.

Problem 3 (Ever Decreasing Stalk Market) The stalk market contains a pattern that a starting price can be as low as 0.4, and then price can drop for at most 6 more consecutive sessions. In ith session, the amount of the drop is $X_i \sim \mathrm{Unif}(0.03, 0.05)$, where X_i's are independent random variables. Denote the price after n more sessions as $P_n = 0.4 + \sum_{i=1}^{n} X_i$. We are interested in analyzing how bad the market can be.

a) Show that $E[P_n] = 0.4 - 0.04n$ and $V(P_n) = 0.01^2/3 \cdot n$.

b) Draw in the same figure the following quantities against n: the mean price $E[P_n]$, the lower and upper 95% confidence bounds $E[P_n] \pm 1.96\sqrt{V(P_n)}$, the worse price $W_n = 0.4 - 0.05n$. Give an intuition explanation why the worse price line is way pessimistic than the lower confidence bound.

c) Use the convolution theorem for independent random variables, draw the PDFs for P_n, $n = 1, 2$, and 3. From this sequence of PDFs, can you guess the shape of the PDF for P_6? Can you link your guess to any theorem?

References

[1] Nintendo, "Animal crossing: New horizons," Mar. 2020. [Online]. Available: https://www.animal-crossing.com/new-horizons/

[2] /u/Edricus, "Breaking down the stalk market," Apr. 2020. [Online]. Available: https://docs.google.com/document/d/1bSVNpOnH_dKxkAGr718-iqh8s8Z0qQ54L-0mD-lbrXo

[3] Ninji, "AC:NH turnip price calculator," Mar. 2020. [Online]. Available: https://gist.github.com/Treeki/85be14d297c80c8b3c0a76375743325b

(First version: 4/29/2020. If you find any typo/mistake, please email chauwai.wong at ncsu dot edu)