ECE2521: Analysis of Stochastic Processes

Homework 3: Topics 5 and 6 (Random Variables and Distributions)

Submission by: Seth So

Professor Dr. Azime Can-Cimino

Submitted: September 23, 2020

Assigned Problems (11)

Textbook Problems (3.1 - 3.10): 3.2, 3.9, 3.12, 3.17, 3.28, 3.31, 3.42, 3.46, 3.57, and 3.91
Application Problems (3.11): MATLAB: Distribution of Coin Flips

Textbook: (3.1 – 3.7)

3.2. A die is tossed and the random variable X is defined as the number of full pairs of dots in the face showing up.

(a) Describe the underlying space S of this random experiment and specify the probabilities of its elementary events.

(b) Show the mapping from S to S_X, the range of X.

(c) Find the probabilities for the various values of X.

(d) Repeat parts a, b, and c, if Y is the number of full or partial pairs of dots in the face showing up.

(e) Explain why $P[X = 0]$ and $P[Y = 0]$ are not equal.

3.9. A coin is tossed n times. Let the random variable Y be the difference between the number of heads and the number of tails in the n tosses of a coin. Assume $P[\text{heads}] = p$.

(a) Describe the sample space of S.

(b) Find the probability of the event $\{Y = 0\}$.

(c) Find the probabilities for the other values of Y.

3.12. Consider an information source that produces binary pairs that we designate as $S_X = \{1, 2, 3, 4\}$. Find and plot the pmf in the following cases:

(a) $p_k = p_1/k$ for all k in S_X.

(b) $p_{k+1} = p_k/2$ for $k = 2, 3, 4$.

(c) $p_{k+1} = p_k/2^k$ for $k = 2, 3, 4$.

(d) Can the random variables in parts a, b, and c be extended to take on values in the set $\{1, 2, \ldots\}$? If yes, specify the pmf of the resulting random variables. If no, explain why not.

3.17. A modem transmits a +2 voltage signal into a channel. The channel adds to this signal a noise term that is drawn from the set $\{0, -1, -2, -3\}$ with respective probabilities $\{4/10, 3/10, 2/10, 1/10\}$.

(a) Find the pmf of the output Y of the channel.

(b) What is the probability that the output of the channel is equal to the input of the channel?

(c) What is the probability that the output of the channel is positive?

3.28. Find the expected value and variance of the modem signal in Problem 3.17.

3.31. **(a)** Suppose a fair coin is tossed n times. Each coin toss costs d dollars and the reward in obtaining X heads is $aX^2 + bX$. Find the expected value of the net reward.

(b) Suppose that the reward in obtaining X heads is a^X, where $a > 0$. Find the expected value of the reward.

3.42. Find $E[Y]$ and $\text{VAR}[Y]$ for the difference between the number of heads and tails in n tosses in Problem 3.9. *Hint:* Condition on the number of heads.

Textbook: (3.8–3.10) + Application (3.11)

3.46. Heat must be removed from a system according to how fast it is generated. Suppose the system has eight components each of which is active with probability 0.25, independently of the others. The design of the heat removal system requires finding the probabilities of the following events:

(a) None of the systems is active.

(b) Exactly one is active.

(c) More than four are active.

(d) More than two and fewer than six are active.

3.57. A Christmas fruitcake has Poisson-distributed independent numbers of sultana raisins, iridescent red cherry bits, and radioactive green cherry bits with respective averages 48, 24, and 12 bits per cake. Suppose you politely accept 1/12 of a slice of the cake.

(a) What is the probability that you get lucky and get no green bits in your slice?

(b) What is the probability that you get really lucky and get no green bits and two or fewer red bits in your slice?

(c) What is the probability that you get extremely lucky and get no green or red bits and more than five raisins in your slice?

3.91. The number X of photons counted by a receiver in an optical communication system is a Poisson random variable with rate λ_1 when a signal is present and a Poisson random variable with rate $\lambda_0 < \lambda_1$ when a signal is absent. Suppose that a signal is present with probability p.

(a) Find $P[\text{signal present} \mid X = k]$ and $P[\text{signal absent} \mid X = k]$.

(b) The receiver uses the following decision rule:

If $P[\text{signal present} \mid X = k] > P[\text{signal absent} \mid X = k]$, decide signal present; otherwise, decide signal absent.

Show that this decision rule leads to the following threshold rule:

If $X > T$, decide signal present; otherwise, decide signal absent.

(c) What is the probability of error for the above decision rule?

Problem 3.11: Write a MATLAB script to produce a random variable which follows a Binomial distribution for arbitrary parameters n, and p with n as the number of the trials and p as the probability of observing a head in one coin toss. Estimate the probability mass function (PMF) of this random variable using relative frequency approach and compare it with the actual PMF by plotting the estimated and actual PMFs for $n = 10$ and $p = 0.3$. Submit a printout of your script together with the comparison plot. (Hint: Use the MATLAB function *nchoosek* (n,k) to compute the binomial coefficients in the actual PMF.)

Textbook Problems 1 through 10

3.1 (3.2):

A die is tossed with X := of full pairs of dots showing:

(a) Underlying space S:

$$S = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$$
$$P(S_i) = [\frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}, \frac{1}{6}]$$

(b) S $\rightarrow$ S$_X$:

$$S = \{1,\ 2,\ 3,\ 4,\ 5,\ 6\}$$
$$S_X = \{0,\ 1,\ 1,\ 2,\ 2,\ 3\}$$

(c) Probability of element in S$_X$:

$$X = \{0,\ 1,\ 2,\ 3\}$$
$$P_X = [\frac{1}{6}, \frac{1}{3}, \frac{1}{3}, \frac{1}{6}]$$

(d) Full or partial pairs = X$_{full}$ + Y$_{partial}$

$$\begin{aligned} S_X &= \{0,\ 1,\ 1,\ 2,\ 2,\ 3\} \\ S_Y &= \{1,\ 0,\ 1,\ 0,\ 1,\ 0\} \\ S_{X+Y} &= \{1,\ 1,\ 2,\ 2,\ 3,\ 3\} \\ P_{X+Y} &= [\frac{1}{3}, \frac{1}{3}, \frac{1}{3}] \text{ for outcomes } \mathrm{S}_{X+Y} = \{1,2,3\}, \text{ respectively} \end{aligned}$$

(e) The probabilites are different since X{0} $\rightarrow$ S{1}, while Y{0} $\rightarrow$ S{1,2}

3.2 (3.9):

A coin is tossed n times, let Y := (#heads)-(#tails):

(a) S$_Y$ = {-n to n}, for {no heads, all heads}

(b) Y = 0 $\rightarrow$ #heads = #tails:

$$P[heads = \frac{flips}{2}] = \boxed{\binom{n}{\frac{n}{2}} * p^{\frac{n}{2}} * (1-p)^{\frac{n}{2}},\ for\ n_{even}}$$

(c) For any given Y:

$$\begin{aligned} P[Y|k_{heads}] &= P[heads] - P[tails] \\ &= (\binom{n}{k} * p^k * (1-p)^{n-k}) - (1 - \binom{n}{k} * p^k * (1-p)^{n-k}) \\ &= \boxed{2(\binom{n}{k} * p^k * (1-p)^{n-k}) - 1} \end{aligned}$$

3.3 (3.12):

Given $S_X = \{1, 2, 3, 4\}$:

(a) PMF of $p_k = \frac{p_1}{k} \forall\, k \in S_X$:

$$\begin{aligned} 1 &= p_1 + p_2 + p_3 + p_4 \\ &= p_1 + \frac{p_1}{2} + \frac{p_1}{3} + \frac{p_1}{4} \\ &= \boxed{p_1 * \sum_{k=0}^{3} \frac{1}{k+1}, \quad \therefore \quad p_1 = \frac{12}{25}} \end{aligned}$$

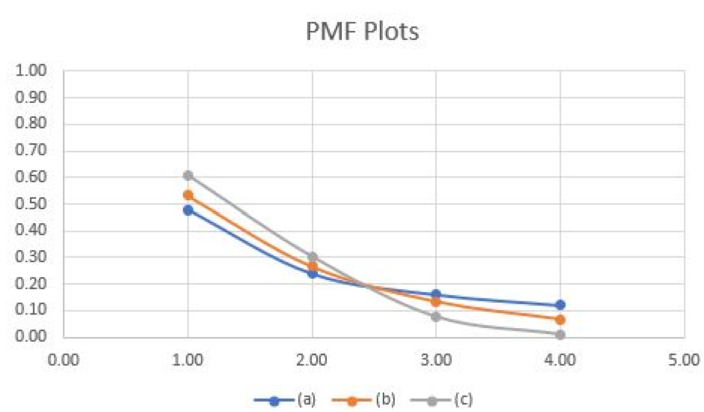
(b) PMF of $p_{k+1} = \frac{p_k}{2}\ for\ k = 2, 3, 4$:

$$\begin{aligned} 1 &= p_1 + p_2 + p_3 + p_4 \\ &= p_1 + \frac{p_1}{2} + \frac{p_1}{4} + \frac{p_1}{8} \\ &= \boxed{p_1 * \sum_{k=0}^{3} 2^{-k}, \quad \therefore \quad p_1 = \frac{8}{15}} \end{aligned}$$

(c) PMF of $p_{k+1} = \frac{p_k}{2^k}\ for\ k = 2, 3, 4$:

$$\begin{aligned} 1 &= p_1 + p_2 + p_3 + p_4 \\ &= p_1 + \frac{p_1}{2} + \frac{p_1}{8} + \frac{p_1}{64} \\ &= \boxed{p_1 * \sum_{k=0}^{3} \frac{1}{2^{k*\frac{k-1}{2}}}, \quad \therefore \quad p_1 = \frac{64}{105}} \end{aligned}$$

(d) Below are the PMFs for the previous parts, populated by solving for and plotting p_{1-4} for each part in Excel:



For the random variables to be able to take on S = {1, 2, 3, ...}, they must converge under the natural set. Using the series from each part, it is found that (a) does not converge, while (b) and (c) do converge.

3.4 (3.17):

Given +2V + noise values {0, -1, -2, -3} with respective probabilities P[$\frac{4}{10}, \frac{3}{10}, \frac{2}{10}, \frac{1}{10}$]

(a) PMF = {2, 1, 0 ,-1}, corresponding to elements of P above, respectively

(b) P[in = out] = P[+0V] = $\boxed{\frac{2}{5}}$

(c) P[out > 0] = $\frac{4}{10} + \frac{3}{10} = \boxed{\frac{7}{10}}$

3.5 (3.28):

Given information from 3.17 to define r.v. X:

(a) $E[X] = \sum_x x * P[x] = (2 * \frac{4}{10}) + (1 * \frac{3}{10}) + (0 * \frac{2}{10}) + (-1 * \frac{1}{10}) = \boxed{1}$

(b) $Var[X] = E[X^2] - E[X]^2 = (4 * \frac{4}{10}) + (1 * \frac{3}{10}) + (0 * \frac{2}{10}) + (1 * \frac{1}{10}) - 1 = \boxed{1}$

3.6 (3.31):

(a) n tosses, d \$ cost each, X heads, aX^2 + bX reward

$$E[X] \;=\; Reward - Cost \;=\; E[aX^2 + bX - nd] \;=\; aE[X^2] + bE[X] - nd$$

Solving for E[X^2] and E[X] with known Binomial Distribution properties:

$$\begin{aligned} E[X^2] &= Var[X] + E[X]^2 \\ &= np(1-p) + (np)^2 \\ &= n(\frac{1}{2})(\frac{1}{2}) + n^2(\frac{1}{4}) \\ &= \frac{n}{4} + \frac{n^2}{4} \end{aligned} \qquad \begin{aligned} E[X] &= \sum_k kP[k] \\ &= np \\ &= n(\frac{1}{2}) \\ &= \frac{n}{2} \end{aligned}$$

$$\therefore NetProfit = \boxed{a(\frac{n^2}{4} + \frac{n}{4}) + b(\frac{n}{2}) - nd}$$

(b) n tosses, d \$ cost each, X heads, a^X reward

$$\begin{aligned} E[X] &= \sum_{k=0}^{n} a^k \binom{n}{k} \left(\frac{1}{2}\right)^n \\ &= \sum_{k=0}^{n} a^k \binom{n}{k} \left(\frac{1}{2}\right)^{n-k} \left(\frac{1}{2}\right)^k \\ &= \sum_{k=0}^{n} \binom{n}{k} \left(\frac{1}{2}\right)^{n-k} \left(\frac{a}{2}\right)^k \\ &= \boxed{(\frac{1}{2} + \frac{a}{2})^n} \end{aligned}$$

3.7 (3.42):

Find E[Y] and Var[Y] from question 3.2 (book 3.9):

Y = #heads - #tails → for n flips and k heads, there are (n-k) tails.
Therefore Y = k - (n-k) = **2k - n**.

(a) E[Y] is then:

$$
\begin{aligned}
E[Y] &= \sum_{k=0}^{n} (2k-n) \binom{n}{k} p^k (1-p)^{n-k} \\
&= 2\sum_{k=0}^{n} (k) \binom{n}{k} p^k (1-p)^{n-k} - n \sum_{k=0}^{n} \binom{n}{k} p^k (1-p)^{n-k} \\
&= 2E[K] - n * 1 \\
&= 2np - n \\
&= \boxed{n(2p-1)}
\end{aligned}
$$

(b) E[Y] can be used to solve for Var[Y]:

$$
\begin{aligned}
Var[Y] &= E[Y^2] - E[Y]^2 \\
&= \left(\sum_{k=0}^{n} \left(4k^2 - 4kn + 2k^2\right) \binom{n}{k} p^k (1-p)^{n-k} \right) - (2np-n)^2 \\
&= \left(4E[k^2] - 4nE[k] + n^2\right) - \left(4n^2p^2 - 4n^2p + n^2\right) \\
&= \left(4Var[k] + 4E[k]^2 - 4nE[x]\right) - \left(4n^2p^2 - 4n^2p + n^2\right) \\
&= \left(4np(1-p) + 4n^2p^2 - 4n^2p + n^2\right) - \left(4n^2p^2 - 4n^2p + n^2\right) \\
&= \boxed{4np(1-p)}
\end{aligned}
$$

3.8 (3.46):

8 components with independent 0.25 chance of being active, find:

(a) Active = 0: $\sum_{k=0}^{0} \binom{8}{k} * 0.25^k * (0.75)^{8-k} = \boxed{0.1001}$

(b) Active = 1: $\sum_{k=1}^{1} \binom{8}{k} * 0.25^k * (0.75)^{8-k} = \boxed{0.267}$

(c) Active > 4: $\sum_{k=5}^{8} \binom{8}{k} * 0.25^k * (0.75)^{8-k} = \boxed{0.0273}$

(d) 2 < Active < 6: $\sum_{k=3}^{5} \binom{8}{k} * 0.25^k * (0.75)^{8-k} = \boxed{0.317}$

These and subsequent series were evaluated using MATLAB *symsum* function

3.9 (3.57):

Givens: $\lambda = \frac{1}{12}$ region, $k_s = 48$, $k_r = 24$, $k_g = 12$, with rate α = k*λ bits/slice.
If bit placement is independent, find:

(a) P[=0green] $= \sum_{k=0}^{0} \frac{(k_g*\lambda)^k}{k!} e^{-k_g*\lambda} = e^{-1} = \boxed{0.368}$

(b) P[=0green] & P[< 3red] $= e^{-1} * \sum_{k=0}^{2} \frac{(k_r*\lambda)^k}{k!} e^{-k_r*\lambda} = \boxed{0.249}$

(c) P[=0green] & P[=0red] & P[> 5sultana] $= e^{-1} * \sum_{k=0}^{0} \frac{(k_r*\lambda)^k}{k!} e^{-k_r*\lambda} * \sum_{k=6}^{48} \frac{(k_s*\lambda)^k}{k!} e^{-k_s*\lambda} = \boxed{0.0107}$

3.10 (3.91):

X photons, λ_1 with signal & λ_0 without signal. P[sig] = p. Find:

(a) Conditional probabilites:

$$\begin{aligned}
P[sig|X=k] &= \frac{P[(X=k)\cap sig]}{P[X=k|sig]P[sig]+P[X=k|nosig]P[nosig]} \\
&= \frac{P[(X=k)P[sig]}{P[X=k|sig]P[sig]+P[X=k|nosig]P[nosig]} \\
&= \boxed{\frac{\frac{\lambda_1^k}{k!}e^{-\lambda_1}*p}{\frac{\lambda_1^k}{k!}e^{-\lambda_1}*p+\frac{\lambda_0^k}{k!}e^{-\lambda_0}*(1-p)}}
\end{aligned}$$

Similarly:

$$\begin{aligned}
P[nosig|X=k] &= \frac{P[(X=k)\cap nosig]}{P[X=k|sig]P[sig]+P[X=k|nosig]P[nosig]} \\
&= \frac{P[(X=k)P[nosig]}{P[X=k|sig]P[sig]+P[X=k|nosig]P[nosig]} \\
&= \boxed{\frac{\frac{\lambda_0^k}{k!}e^{-\lambda_0}*(1-p)}{\frac{\lambda_1^k}{k!}e^{-\lambda_1}*p+\frac{\lambda_0^k}{k!}e^{-\lambda_0}*(1-p)}}
\end{aligned}$$

(b) Find threshold T such that X > T results in deciding if there is a signal:

$$\begin{aligned}
P[sig|X=k] &> P[nosig|X=k] \\
\frac{\frac{\lambda_1^k}{k!}e^{-\lambda_1}*p}{\frac{\lambda_1^k}{k!}e^{-\lambda_1}*p+\frac{\lambda_0^k}{k!}e^{-\lambda_0}*(1-p)} &> \frac{\frac{\lambda_0^k}{k!}e^{-\lambda_0}*p}{\frac{\lambda_1^k}{k!}e^{-\lambda_1}*p+\frac{\lambda_0^k}{k!}e^{-\lambda_0}*(1-p)} \\
\lambda_1^k e^{-\lambda_1}*p &> \lambda_0^k e^{-\lambda_0}*(1-p) \\
\frac{\lambda_1}{\lambda_0}^k &> e^{\lambda_1-\lambda_0}\frac{(1-p)}{p} \\
k &> \boxed{\frac{\lambda_1-\lambda_0+ln\left(\frac{1-p}{p}\right)}{ln\left(\frac{\lambda_1}{\lambda_0}\right)}}
\end{aligned}$$

(c) Error with above rule:

$$\begin{aligned}
P_{error} &= P[X<T|sig]P[sig]+P[X\geq T|nosig]P[nosig] \\
&= \boxed{\left(\sum_{k=0}^{<T}\frac{\lambda_1^k}{k!}e^{-\lambda_1}\right)*p+\left(\sum_{k=T}^{\infty}\frac{\lambda_0^k}{k!}e^{-\lambda_0}\right)*(1-p)}
\end{aligned}$$

Application Problem 11

3.11 (Coin Flip Probability):

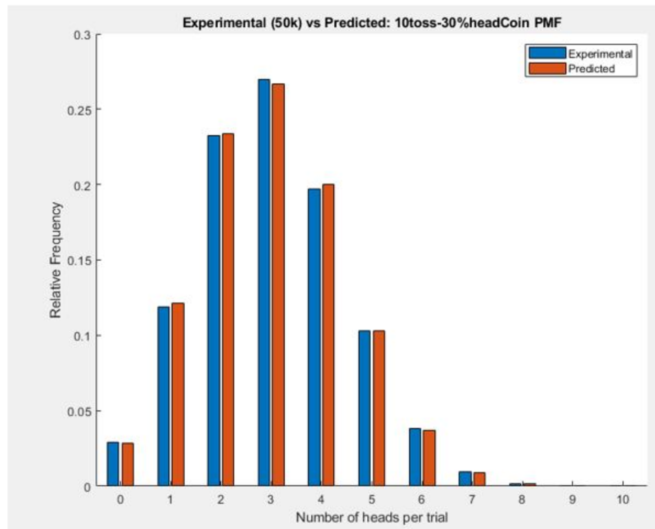
Problem done entirely in MATLAB. Note that the 'n' and 'p' mentioned in the problem are 'numToss' and 'probHead' in the script, respectively:

Editor - A:\Matlab\ECE2521\hw3_11_coinPMF.m

hw3_11_coinPMF.m

```matlab
%Seth So 9/22/20
%ECE2521 hw3q11

clc
clear all
close all

%note currently the table only generates for <10 due to matrix agreements
numToss = 10;
probHead = 0.3;
trials = 50000;

[expRelFreq] = predictedCoinPMF(trials, numToss, probHead);
[predRelFreq] = actualCoinPMF(numToss, probHead);

errorPercent = (predRelFreq-expRelFreq)./predRelFreq;
outcomes = [0:numToss]';
expCount = expRelFreq*trials;
predCount = predRelFreq*trials;

table(outcomes, expRelFreq,predRelFreq,expCount,predCount,errorPercent)

comparePlot = bar(outcomes,[expRelFreq' ; predRelFreq']);
xlabel('Number of heads per trial')
ylabel('Relative Frequency')
title('Experimental (50k) vs Predicted: 10toss-30%headCoin PMF')
legend ('Experimental', 'Predicted')

function [expRelFreq] = predictedCoinPMF(trials, numToss,prob)
    perTen = prob*10;
    probSim = [zeros(1,10-perTen) ones(1,perTen)];
    resultTotal = [];
    for i = 1:trials
        resultTrial = 0;
        for j = 1:numToss
            resultTrial = resultTrial + probSim(randperm(10,1));
        end
        resultTotal = [resultTotal resultTrial];
    end
    [counts,items] = groupcounts(resultTotal');
    if ismember(10,items) == false
        counts = [counts ; 0];
    end
    expRelFreq = counts/trials;
    hold on
end

function [predRelFreq] = actualCoinPMF(numToss, prob)
    syms k
    predRelFreq = [];
    for i = 0:numToss
        outcomeProb = nchoosek(numToss,i)*prob^(i)*(1-prob)^(numToss-i);
        predRelFreq = [predRelFreq outcomeProb];
    end
    predRelFreq = predRelFreq';
    hold on
end
```

Script, Generated Figure, and Command line Output



outcomes	expRelFreq	predRelFreq	expCount	predCount	errorPercent
0	0.02678	0.028248	1339	1412.4	0.051952
1	0.12054	0.12106	6027	6053	0.0043021
2	0.23352	0.23347	11676	11674	-0.00019514
3	0.26836	0.26683	13418	13341	-0.0057418
4	0.20064	0.20012	10032	10006	-0.0025937
5	0.10194	0.10292	5097	5146	0.0095157
6	0.03716	0.036757	1858	1837.8	-0.010966
7	0.00966	0.0090017	483	450.08	-0.073132
8	0.00128	0.0014467	64	72.335	0.11523
9	0.00012	0.00013778	6	6.889	0.12905
10	0	5.9049e-06	0	0.29524	1

>>