Solutions to Homework 2 - Probability Review

1 Bernoulli, binomial, Poisson and normal distributions.

A Binomial distribution. Since X_n is a binomial RV with parameters (n, p), it can be written as

$$X_n = \sum_{i=1}^{n} B_i \tag{1}$$

where $B_1, \ldots, B_n$ are i.i.d. Bernoulli RVs with parameter p (i.e., $\mathrm{P}(B_i = 1) = p$). From the linearity of the expectation operator, we have from (1)

$$\mathbb{E}[X_n] = \mathbb{E}\left[\sum_{i=1}^{n} B_i\right] = \sum_{i=1}^{n} \mathbb{E}[B_i] = n\mathbb{E}[B_1].$$

The expectation of the Bernoulli RVs is $\mathbb{E}[B_1] = 1 \times \mathrm{P}(B_1 = 1) + 0 \times \mathrm{P}(B_1 = 0) = p$. Hence, it follows that

$$\mathbb{E}[X_n] = n\mathbb{E}[B_1] = np.$$

Now for the variance, since $B_1, \ldots, B_n$ are independent and hence uncorrelated RVs we can write

$$\mathrm{var}[X_n] = \mathrm{var}\left[\sum_{i=1}^{n} B_i\right] = \sum_{i=1}^{n} \mathrm{var}[B_i] = n\mathrm{var}[B_1].$$

To calculate $\mathrm{var}[B_1] = \mathbb{E}\left[(B_1 - \mathbb{E}[B_1])^2\right] = \mathbb{E}\left[B_1^2\right] - \mathbb{E}[B_1]^2$ we should find the second moment $\mathbb{E}\left[B_1^2\right]$ (we already know $\mathbb{E}[B_1]^2 = p^2$). But this is straightforward because for Bernoulli B_1, it holds that $B_1 = B_1^2$. Hence $\mathbb{E}\left[B_1^2\right] = p$ and the sought variance of X_n is

$$\mathrm{var}[X_n] = n\mathrm{var}[B_1] = n(p - p^2) = np(1-p).$$

The following Matlab function calculates the pmf of a binomial RV.

```
% This function returns a vector the same size as u where entries
% are the binomial pmf with parameters n and p, calculated
% at each element of u. This tries to mimic Matlab's
% built-in function f=pdf('bino',u,n,p)
% For how Matlab itself calculates the binomial pdf, see binopdf.m

function f=my_binomial_pmf(u,n,p)
f=zeros(size(u)); %initialization of f
for i=1:length(u)
    if (u(i)>=0) && (u(i)<=n) % since support of binomial(n,p) is 0,1,..,n
        f(i)=nchoosek(n,u(i))*p^u(i)*(1-p)^(n-u(i));
              % pmf expression of binomial RV with parameters (n,p)
    end
end
end
```

Using the function `my_binomial_pmf` the following main script plots the required pmfs and cdfs, for fix $\mathbb{E}[X_n] = 5$ and $n = 6, 10, 20, 50$. Notice that for discrete RVs, we can use the command `stem` for plotting the pmf and `stairs` for the cdf.

```
clear all; close all; clc;

n_vector=[6,10,20,50];
```

```
i=1;
for n=n_vector
    p=5/n;

    figure(1) % pmfs
    subplot(2,2,i);
    % stem(0:n,pdf('bino',0:n,n,p),'.'); % cheat line!
    stem(0:n,my_binomial_pmf(0:n,n,p),'.');
    title(['n=',num2str(n)]); xlabel('x'); ylabel('pmf');
    grid on; axis([0,50,0,0.5]);

    figure(2) % cdfs
    subplot(2,2,i);
    % stairs(0:n,cdf('bino',0:n,n,p),'.'); % cheat line!
    % stairs(0:n,my_binomial_cdf(0:n,n,p),'.'); % see my_binomial_cdf below
    stairs(0:n,cumsum(my_binomial_pmf(0:n,n,p)),'LineWidth',1); % with cumsum
    title(['n=',num2str(n)]); xlabel('x'); ylabel('cdf');
    grid on; axis([0,50,0,1]);

    i=i+1;
end
```

In case you prefer a separate function to calculate the binomial cdf:

```
% function takes vector u, and scalars n and p,
% and returns a vector of the same size as u, where entries
% are the binomial cdf with parameters n and p, calculated
% at the points of elements of u. trying to mimic Matlab's
% built-in function f=cdf('bino',u,n,p)
% This function calls my_binomial_pdf and calculates a cumulative sum
% over possible values up to u(i) for each entry of u.

function F=my_binomial_cdf(u,n,p)
F=zeros(size(u)); %initialization of F
for i=1:length(u)
    F(i)=sum(my_binomial_pdf(0:u(i),n,p));
end
```

The obtained pmf and cdf plots are shown in Figs. 1 and 2.

B Binomial and Poisson distributions. For a Poisson RV X_p with parameter λ, we went through the calculation of $\mathbb{E}[X_P]$ in class (check the lecture slides). A different way of evaluating the expected value is

$$\mathbb{E}[X_P] = \sum_{k=0}^{\infty} k\mathrm{P}(X_p = k) = \sum_{k=0}^{\infty} k \frac{e^{-\lambda}\lambda^k}{k!} = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{k\lambda^{k-1}}{k!}. \tag{2}$$

Recall the Taylor series expansion $e^{\lambda} = \sum_{k=0}^{\infty} \frac{\lambda^k}{k!}$, and from the linearity of the differentiation operator we have

$$\sum_{k=0}^{\infty} \frac{k\lambda^{k-1}}{k!} = \frac{d}{d\lambda}\left(\sum_{k=0}^{\infty} \frac{\lambda^k}{k!}\right) = \frac{d}{d\lambda}\left(e^{\lambda}\right) = e^{\lambda}.$$

Plugging this result back in (2) yields

$$\mathbb{E}[X_P] = \lambda e^{-\lambda} \sum_{k=0}^{\infty} \frac{k\lambda^{k-1}}{k!} = \lambda e^{-\lambda} e^{\lambda} = \lambda.$$

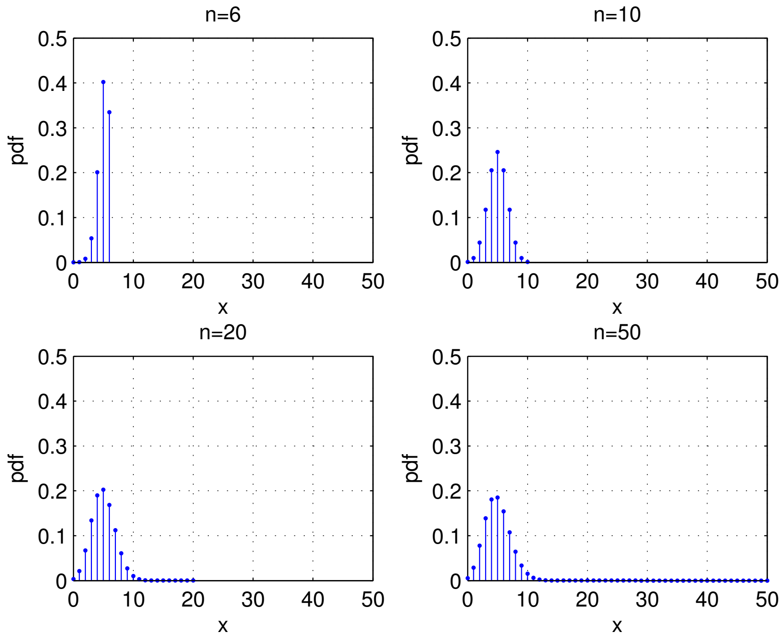

Fig. 1. Binomial pmf for $n = 6, 10, 20, 50$ and $p = 5/n$. (Part A).

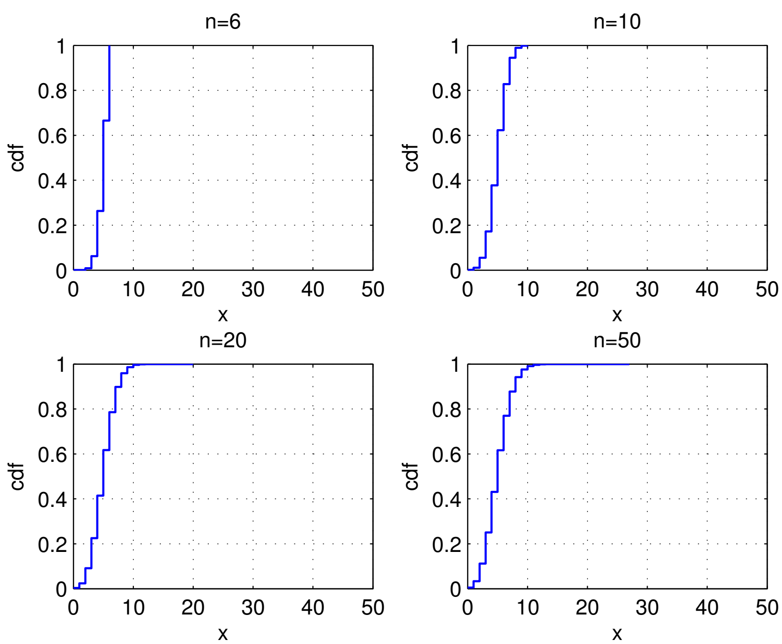

Fig. 2. Binomial cdf for $n = 6, 10, 20, 50$ and $p = 5/n$. (Part A).

The Matlab code to plot the pmf of a Poisson distribution with parameter $\lambda = 5$ follows.

```
clear all; close all; clc;
figure
lambda=5;
```

```
x=0:50;
% stem(x,pdf('poiss',x,lambda),'.'); % cheat line!
my_poisson_pmf=exp(-lambda)*(lambda.^x)./factorial(x);
    % Note the use of "dot" for element-wise operation
    % Thus, my_poisson_pmf is now a vector with the same size as x.
stem(x,my_poisson_pmf,'.');
title(['Poisson distribution with \lambda = ',num2str(lambda)]);
xlabel('x'); ylabel('pmf');
grid on; axis([0,50,0,0.5]);
```

The obtained pmf is depicted in Fig. 3.

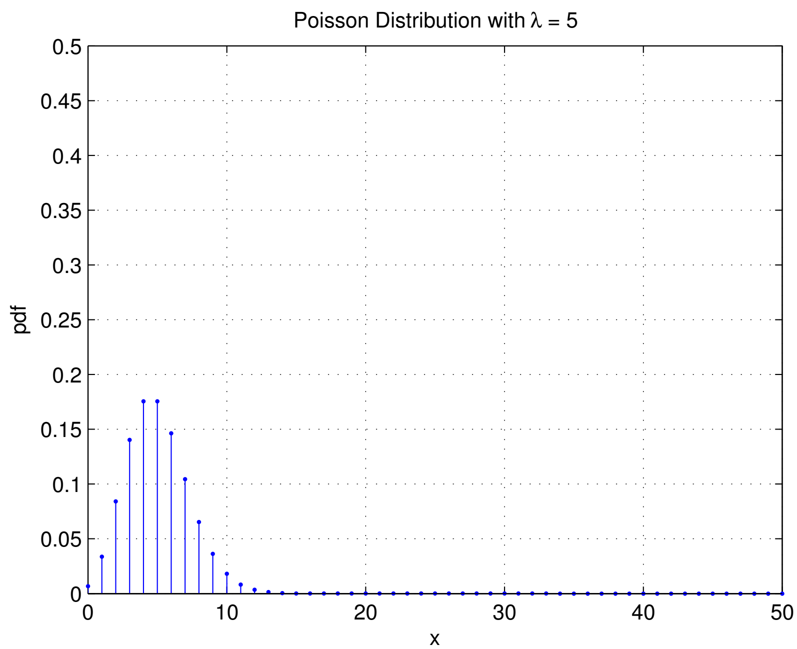

Fig. 3. Pmf of the Poisson distribution with $\lambda = 5$. Note that the support of a Poisson RV are the nonnegative integers. However, the pmf for only the first 50 points is shown. Note the similarity with Fig. 1 for large n.

Now we need to calculate the mean-squared error (MSE) between the binomial and Poisson pmfs. The MSE is defined as $\Delta(X_n, X_P) = \sum_{x=0}^{\infty} \left(\mathrm{P}\left(X_n = x\right) - \mathrm{P}\left(X_P = x\right) \right)^2 \mathrm{P}\left(X_P = x\right)$. To numerically evaluate the MSE, the infinite sum needs to be truncated by e.g., neglecting probabilities smaller than 5×10^{-2}. To identify those small probabilities, the following code reveals that we need not go beyond $x = 8$ in the MSE summation (note from Fig. 3 that the Poisson distribution has a declining tail, which meas that as $x \to \infty$, $\mathrm{P}\left(X_P = x\right) \to 0$).

```
clear all; close all; clc;
x=0:20;
lambda=5;
my_pdf_poisson=exp(-lambda)*lambda.^x./factorial(x);
test=my_pdf_poisson<5e-2 % Entries equal to one indicate small probabilities
```

More precisely, it turns out we should consider only the terms for $x = 2, \ldots, 8$ since $\mathrm{P}\left(X_P = 0\right)$ and $\mathrm{P}\left(X_P = 1\right)$ are also small. Note that by looking at Fig. 3, you can also identify those values for which the probability mass is greater that 5×10^{-2}. The following Matlab script evaluates the MSE, and generates the plot in Fig. 4.

```
close all; clear all; clc;
lambda=5;
```

```
x=2:8;
n_index=0;
n_vector=[6 10 20 50];
my_MSE=zeros(1,4);
for n=n_vector
    n_index=n_index+1;
    my_poisson_pmf=exp(-lambda)*(lambda.^x)./factorial(x);
    my_MSE(1,n_index)=sum((my_binomial_pmf(x,n,lambda/n)...
    -my_poisson_pmf).^2.*my_poisson_pmf);
end
stem(n_vector,my_MSE,'*');
xlabel('n'); ylabel('MSE');
grid on; axis([0,50,0,0.02]);
```

The results reported in Table I show the MSE between the binomial and Poisson distributions for $n = 6, 10, 20, 50$. As n increases, the MSE falls rapidly to zero, indicating the distributions become more and more similar for larger ns. It is critical here that for given λ in the Poisson distribution, $p = \lambda/n$ in the binomial. The decreasing MSE is also apparent from Fig. 4.

TABLE I
MEAN-SQUARED ERROR (MSE) BETWEEN THE PMFS OF A POISSON WITH $\lambda = 5$ AND BINOMIALS WITH PARAMETER $(n, \lambda/n)$ FOR $n = 6, 10, 20, 50$. THE POISSON PROBABILITIES SMALLER THAN 0.05 ARE NEGLECTED.

n	MSE
6	0.01727
10	0.00179
20	0.00027
50	0.00003

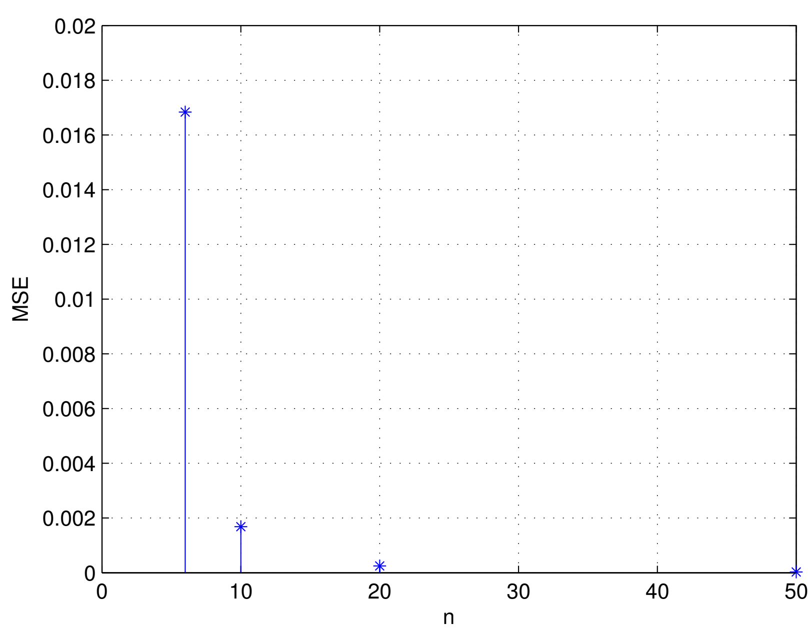

Fig. 4. Mean-Square-Error between pdf of a Poisson with $\lambda = 5$ and Binomials of $(n, \lambda/n)$ for $n = 6, 10, 20, 50$. The probabilities in the Poisson less than 0.05 are neglected. As we see, by increasing n, the MSE rapidly vanishes. (part B)

C Binomial and Poisson distributions again. This is an interesting situation where we can analytically establish what simulations are suggesting. Specifically, as we did in class we will show here that the pmf of a binomial RV X_n with parameters $(n, \lambda/n)$ converges to the pmf of a Poisson RV with parameter λ, as $n \to \infty$. Starting from the expression for the pmf of X_n we have

$$p_{X_n}(x) = \binom{n}{x}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} = \frac{n!}{(n-x)!x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x}. \tag{3}$$

From the definition of factorials, we can simplify

$$\frac{n!}{(n-x)!} = \frac{n(n-1)\dots(n-x+1)(n-x)!}{(n-x)!} = n(n-1)\dots(n-x+1) \tag{4}$$

and rewrite (3) after some reordering of terms to obtain

$$p_{X_n}(x) = \frac{n!}{(n-x)!x!}\left(\frac{\lambda}{n}\right)^x\left(1-\frac{\lambda}{n}\right)^{n-x} = \frac{n(n-1)\dots(n-x+1)}{n^x}\left(\frac{\lambda^x}{x!}\right)\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x}. \tag{5}$$

In order to take the limit as $n \to \infty$ it is useful to recognize that

$$\lim_{n\to\infty}\frac{n(n-1)\dots(n-x+1)}{n^x} = 1, \quad \text{and} \quad \lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^x = 1.$$

In addition to the Taylor series $e^{-\lambda} = \sum_{k=0}^{\infty}\frac{(-\lambda)^k}{k!}$, the function $e^{-\lambda}$ can be equivalently defined as the limit of the sequence

$$\lim_{n\to\infty}\left(1-\frac{\lambda}{n}\right)^n = e^{-\lambda}.$$

Using all these results when taking the limit as $n \to \infty$ in (6) yields the desired result, namely

$$\lim_{n\to\infty} p_{X_n}(x) = \lim_{n\to\infty}\frac{n(n-1)\dots(n-x+1)}{n^x}\left(\frac{\lambda^x}{x!}\right)\frac{\left(1-\frac{\lambda}{n}\right)^n}{\left(1-\frac{\lambda}{n}\right)^x} = \frac{\lambda^x e^{-\lambda}}{x!}. \tag{6}$$

D Binomial and normal distributions. Recall that a binomial RV X_n with parameters (n, p) can be written as

$$X_n = \sum_{i=1}^{n} B_i$$

where $B_1, \dots, B_n$ are i.i.d. Bernoulli RVs with parameter p. We also showed that $\mathbb{E}[B_i] = p$ and $\text{var}[B_i] = p(1-p)$. From the CLT, for sufficiently large n the distribution of

$$Z_n = \frac{\sum_{i=1}^{n} B_i - np}{\sqrt{np(1-p)}} = \frac{X_n - np}{\sqrt{np(1-p)}} \tag{7}$$

is approximately standard normal. From the properties of normal RVs, (7) also implies that

$$X_n = \sqrt{np(1-p)}Z_n + np$$

is approximately normal distributed with mean np and variance $np(1-p)$. In conclusion, for sufficiently large n the cdf of X_n can be approximated as

$$F_{X_n}(x) = \mathrm{P}(X_n \leq x) \approx \frac{1}{\sqrt{2\pi np(1-p)}}\int_{-\infty}^{x} e^{-(u-np)^2/2np(1-p)}\, du.$$

The Matlab code to approximate a binomial cdf with a normal pdf for $p = 0.5$ and $n = 10, 20, 50$ follows:

```
close all; clear all; clc;
n_vector=[10 20 50];
p=0.5;
n_index=0;
for n=n_vector
    n_index=n_index+1;
```

```
    mean_normal=n*p;                        %defining the mean
    std_normal=sqrt(n*p*(1-p));             %defining the standard deviation
    x=0:n;
    subplot(3,1,n_index);                                       %plotting graphs
    stairs(x,[(my_binomial_cdf(x,n,p))' , ...
     normcdf(x,mean_normal,variance_normal)'],'LineWidth',2);
    title(['n=',num2str(n)]); xlabel('x'); ylabel('cdf');
    legend('Binomial','Normal');
    grid on; axis([0,50,0,1]);
end
```

The resulting plots are depicted in Fig 5. As expected from the CLT, by increasing n the normal distribution offers an increasingly accurate approximation of the binomial distribution.

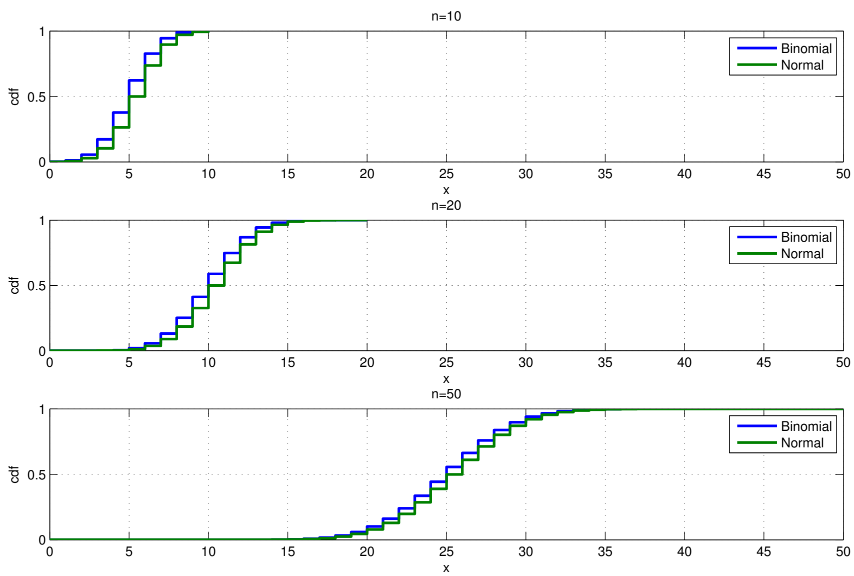

Fig. 5. Cdf of the binomial RV X_n for $n = 10, 20, 50$, and its approximation by a normal cdf. As we can see, by increasing n the normal distribution offers an increasingly accurate approximation of the binomial distribution. This follows from the CLT. (Part D).

E Normal and Poisson approximations. I provide you with a hint for this part: the Poisson limit theorem (also known as law of rare events) is about accumulation of increasingly *improbable* events. In particular, note that for convergence of the distribution of sum of i.i.d. Bernoulli RVs (which is a binomial RV) to a Poisson distribution with mean λ, we needed the success probability in the Bernoulli RV to be $p = \lambda/n$. Accordingly, as $n \to \infty$ this probability goes to zero. On the other hand, for the CLT p is fixed and not necessarily small. Hence, the CLT and Poisson limit theorem are addressing basically different limits. I leave more contemplation on this matter to you!