Probability & Statistics for EECS: Homework #12

Due on May 7, 2023 at 23:59

Name: **Wang Penghao**
Student ID: 2021533138

Problem 1

(a) As for the posterior distribution, we have that is

$$p|X_1 = x_1, X_2 = x_2, ..., X_n = x_n$$

As the prior is Beta(1, 1), then we have that by using the Beta Binomial conjugacy, we get that the posterior with $X_1 = x_1$ is that

$$Beta(1 + x_1, 1 + (1 - x_1))$$

The same, with additional given $X_2 = x_2$, we get that the posterior is that

$$Beta(1 + x_1 + x_2, 1 + (1 - x_1) + (1 - x_2))$$

So, given that $X_1 = x_1, X_2 = x_2, ..., X_n = x_n$ we get that the posterior is that

$$Beta(1 + \sum_{j=1}^{n} x_j, 1 + n - \sum_{j=1}^{n} x_j)$$

So we get that

$$p|X_1 = x_1, X_2 = x_2, ..., X_n = x_n \sim Beta(1 + \sum_{j=1}^{n} x_j, 1 + n - \sum_{j=1}^{n} x_j)$$

From where we see that the parameters of the distribution is only consisted of $\sum_{j=1}^{n} x_j$ so we get that the distribution only depends on the sum of x_j.

(b) Proof of Laplace's law of succession, denote the sum of X_j is that

$$S_n = \sum_{j=1}^{n} X_j,$$

use the LOTP, denote that $f(p|S_n = k)$ is the posterior PDF, we get that

$$P(X_{n+1} = 1|S_n = k) = \int_0^1 P(X_{n+1} = 1|p, S_n = k)f(p|S_n = k)dp$$

then we have that

$$\begin{aligned}
P(X_{n+1} = 1|S_n = k) &= \frac{\Gamma(n+2)}{\Gamma(k+1)\Gamma(n-k+1)} \int_0^1 pp^k(1-p)^{n-k}dp \\
&= \frac{\Gamma(n+2)}{\Gamma(k+1)\Gamma(n-k+1)} \frac{\Gamma(k+2)\Gamma(n-k+1)}{\Gamma(n+3)} \\
&= \frac{k+1}{n+2}.
\end{aligned}$$

(c) From the perspective of Beta-Binomial conjugacy, we have that as the Laplace's Law of Succession states that if we observe an event N times, of which k times are successful, then the probability of success in the next event is $\frac{k+1}{N+2}$. This can be presented as Beta distribution. If we assume a prior probability of beta(a,b), then after observing k successes and (N-k) failures, the posterior probability is beta(a+k,b+(N-k)). The Beta distribution has conjugacy with the binomial distribution, which means that we can use the Beta distribution to represent the posterior distribution of the binomial distribution. Therefore, Laplace's Law of Succession can be reinterpreted as a Bayesian inference process on the binomial distribution using the Beta-Binomial conjugacy relationship.

Problem 2

(a) By using LOTUS, as we have that $p \sim Beta(a, b)$ we have that $E(p^2(1-p)^2)$ can be written as

$$
\begin{aligned}
E(p^2(1-p)^2) &= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 p^2(1-p)^2 p^{a-1}(1-p)^{b-1} dp \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \int_0^1 p^{a+1}(1-p)^{b+1} dp \\
&= \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \frac{\Gamma(a+2)\Gamma(b+2)}{\Gamma(a+b+4)}
\end{aligned}
$$

With the property of Gamma function, we have that

$$\Gamma(a+2) = (a+1)\Gamma(a+1) = (a+1)a\Gamma(a),$$

and that

$$\Gamma(a+b+4) = (a+b+3)\Gamma(a+b+3) = (a+b+3)(a+b+2)(a+b+1)(a+b)\Gamma(a+b).$$

So we get that

$$E(p^2(1-p)^2) = \frac{(a+1)a(b+1)b}{(a+b+3)(a+b+2)(a+b+1)(a+b)}$$

(b) The posterior distribution only depends on the fact that A won exactly 6 of the 10 games on record. As for a sequence event, we need to update the distribution, firstly, we get Beta(a, b), where a = 1, b = 1, then if a A wins occurs, the a += 1, if a B wins occurs, the b += 1. From the update process, we get that the order of outcomes do not effect the final parameters of Beta distribution, so we get that it is the fact that A won exactly 6 of 10 games on record.

(c) As the input event is that AAABBAABAB, where A wins 6 times, and B wins 4 times. So we get that the final posterior distribution for p given historical data is

$$p|historicaldata \sim Beta(6+1, 4+1)$$

which is

$$p|historicaldata \sim Beta(7, 5)$$

(d) As for the indicator of A, we have that they are conditionally independent given p, however, without given p, they are not independent. Thus, given p, the indicators are uncorrelated, without given p, the indicators are postively correlated as if A wins at the first match, the p will increased, as the A is more believed to win.

(e) As for the probability that the match is not yet decided when going into the fifth game There must be 2 wins for A and 2 wins for B, as in (c), we get the posterior distribution is Beta(7, 5), so we get that the probability is

$$\binom{4}{2} p^2(1-p)^2 = 6 \frac{(7+1)7(5+1)5}{(7+5+3)(7+5+2)(7+5+1)(7+5)} = \frac{4}{13}.$$

Problem 3

1. As for the joint PDF of $U_{(1)}, U_{(2)}, ..., U_{(n)}$, as we have that they have the order relationship, we then get that
$$0 < u_{(1)} < ... < u_{(n)} < 1,$$
denote that the joint PDF is $f_{U_{(1)},...,U_{(n)}}$, we then get that as there are $n!$ possible orderings of the $U_{(i)}$,
$$f_{U_{(1)},...,U_{(n)}} = n! \prod_{i=1}^{n} f_{U_{(i)}}(u_{(i)})$$
Then as all the $U_{(i)}$ are i.i.d. uniform distribution, we get that
$$f_{U_{(1)},...,U_{(n)}} = n! \prod_{i=1}^{n} 1 = n!$$

2. As for the joint PDF of $U_{(j)}, U_{(k)}$, where $1 \leq j < k \leq n$, then we denote that the joint PDF is $f_{U_{(j)},U_{(k)}}$, we then firstly calculate the number of permutations of $U_{(j)}, U_{(k)}$, we firstly need to choose j - 1 from n numbers, then 1 for j, the same for others. We get that the number of permutations is
$$\binom{n}{1}\binom{n-1}{1}\binom{n-2}{j-1}\binom{n-j-1}{k-j-1}\binom{n-k}{n-k} = \frac{n!}{(j-1)!(k-j-1)!(n-k)!}$$
The same as (a), we get that CDF for Unif distribution for $U_{(k)}$ is that $u_{(k)}$. With the theorem of PDF of order statistic, we then get that the joint PDF is
$$\begin{aligned} f_{U_{(j)},U_{(k)}} &= \frac{n!}{(j-1)!(k-j-1)!(n-k)!} f(u_{(j)})(u_{(j)})^{j-1}(u_{(k)} - u_{(j)})^{k-j-1} f(u_{(k)})(1-u_{(k)})^{n-k} \\ &= \frac{n!}{(j-1)!(k-j-1)!(n-k)!}(u_{(j)})^{j-1}(u_{(k)} - u_{(j)})^{k-j-1}(1-u_{(k)})^{n-k} \end{aligned}$$

3. As we have that $X \sim Bin(n,p)$ and that $B \sim Beta(j, n-j+1)$, we then denote that
$$U_1, U_2, ..., U_n i.i.d. Unif(0,1),$$
which is a sequence of trails, we then denote that X is the number of success, where $X \sim Bin(n,p)$. We then get that $X \geq j$, is the same as the event that $U_{(j)} \leq p$, which is that the j th one in order is at the left of p. Then with the theorem of order statistic, we get that
$$f_{U_{(j)}}(x) = n\binom{n-1}{j-1} x^{j-1}(1-x)^{n-j}$$
Then we get that the PDF of Beta(j, n - j + 1) distribution is
$$f(x) = n\binom{n-1}{j-1} x^{j-1}(1-x)^{n-j}$$
which is the same. So we get that $U_{(j)}$ is the same as Beta(j, n - j + 1). So we get that $P(X \geq j) = P(B \leq p)$

4. As no calculus is allowed, we then have that as for $X \sim Bin(n,p)$ and $B \sim Beta(j, n-j+1)$, we get that $P(X \geq j) = P(B \leq p)$ So we get as for the equation, we have that $P(U_{(j)} \leq x)$ is that
$$\int_0^x \frac{n!}{(j-1)!(n-j)!} t_{j-1}(1-t)^{n-j} dt$$

Problem 3 continued on next page...

As for the $Bin(n,p)$, the probability of $P(X \geq j)$ is

$$P(X \geq j) = \sum_{k=j}^{n} \binom{n}{k} x^k (1-x)^{n-k}$$

As we get in (c), we get that $P(X \geq j) = P(U_{(j)} \leq p)$, so we get that

$$\int_0^x \frac{n!}{(j-1)!(n-j)!} t_{j-1} (1-t)^{n-j} dt = \sum_{k=j}^{n} \binom{n}{k} x^k (1-x)^{n-k}$$

Problem 4

As for the Poisson-Gamma Duality, we prove it based on the model of Poisson process.
Denote that $X_1, X_2, ..., X_{k+1}$ i.i.d. Expo(1), then we have that with the property, we get that

$$Z = X_1 + X_2 + ... + X_{k+1} \sim Gamma(k+1, 1)$$

Then, we based on the Poisson process model, denote that N_t is the number of arrivals within time t, then we have that

$$N_t \sim Pois(\lambda t)$$

which is

$$P(N_t = n) = \frac{e^{-\lambda t}(\lambda t)^n}{n!}, n \geq 0$$

So we get that

$$P(Z > \lambda) = P(N_t \leq k) = \sum_{n=0}^{k} P(N_\lambda = n) = P(X \leq k)$$

So we based on the Poisson process get that Poisson-Gamma Duality

$$P(X \leq k) = P(Z > \lambda)$$

HW12-sol-p05

May 7, 2023

0.0.1 Problem 05 Programming Assignment

(a) Use the Acceptance-Rejection Method to obtain the samples from distribution Beta(2, 4). You need to plot the pictures of both histogram and the theoretical PDF

We need to calculate the constant in Beta distribution, as it is Beta(2, 4), which is

$$\int_0^1 \frac{x(1-x)^3}{c}dx = \frac{1}{c}\int_0^1 x - 3x^2 + 3x^3 - x^4 dx = \frac{1}{c}\frac{1}{20} = 1$$

So we get that

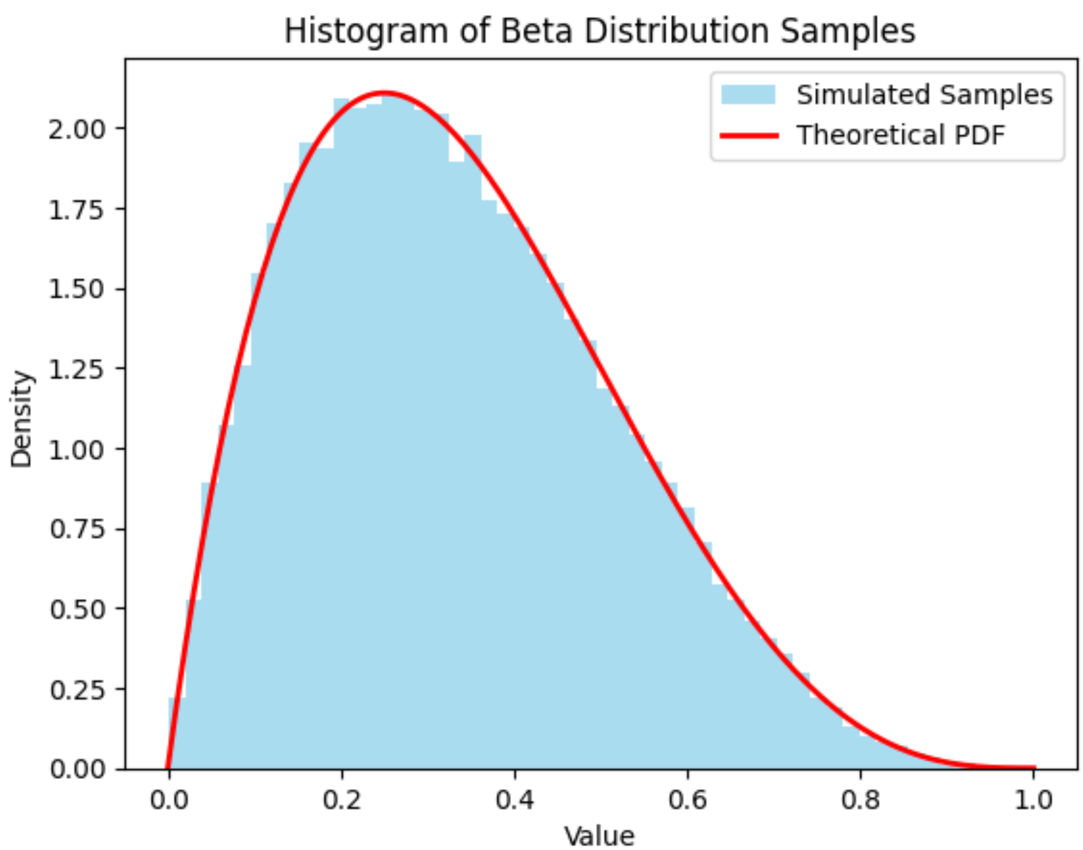
$$c = \frac{1}{20}$$

, then the PDF can be written as

$$f = 20x(1-x)^3$$

[ ]:
```python
import numpy as np
import matplotlib.pyplot as plt

n_sim = 100000
a, b = 2, 4
sample_list = []

def beta_2_4_pdf(x):
    return 20 * x * (1 - x)**3

def Uniform_pdf(x, min, max):
    return 1 / (max - min)

# get the constant c where the PDF of uniform distribution is a constant
c = beta_2_4_pdf((a - 1) / (a + b - 2)) / 1

for i in range(n_sim):
    # g is Uniform distribution
    x = np.random.uniform(0, 1)
    # U is Uniform distribution
    U = np.random.uniform(0, 1)
    if U < beta_2_4_pdf(x) / (c * Uniform_pdf(x, 0, 1)):
        sample_list.append(x)
```

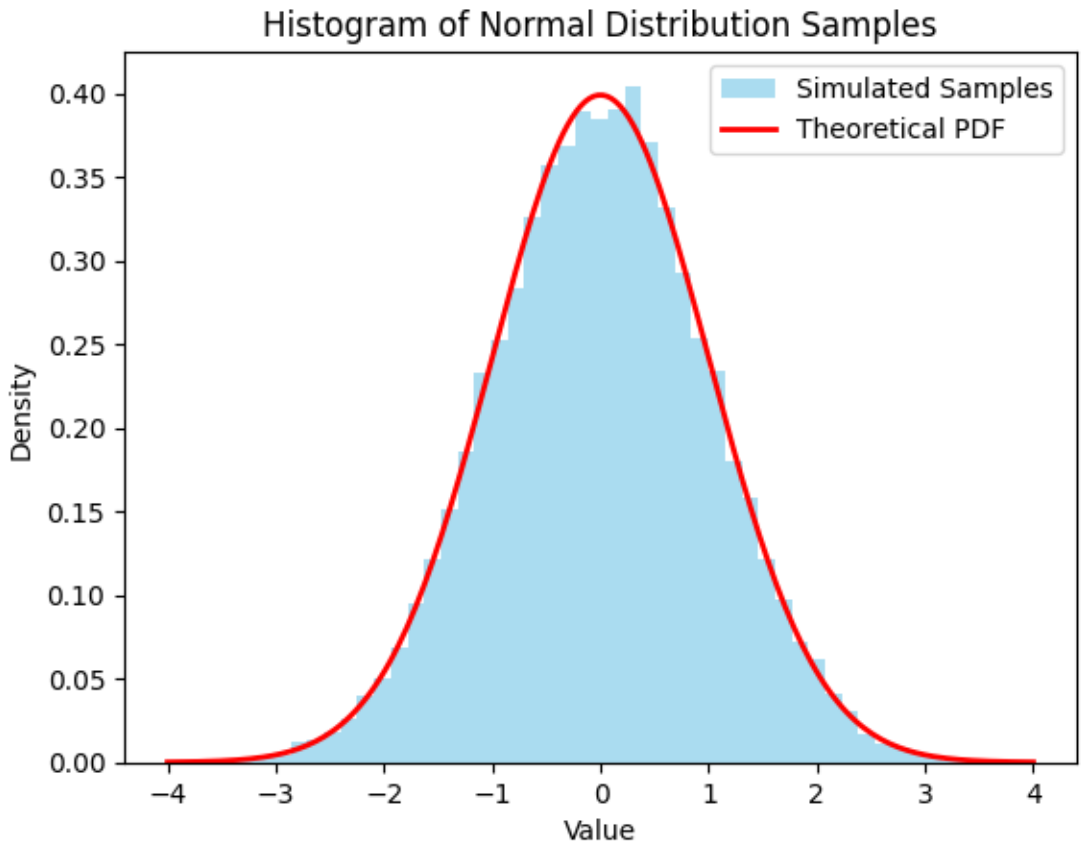
```
sample_array = np.array(sample_list)

x = np.linspace(0, 1, 200)
pdf = beta_2_4_pdf(x)

plt.hist(sample_array, bins=50, density=True, color="skyblue", alpha=0.7,␣
 ↪label="Simulated Samples")
plt.plot(x, pdf, color="red", linewidth=2, label="Theoretical PDF")

plt.legend()
plt.xlabel("Value")
plt.ylabel("Density")
plt.title("Histogram of Beta Distribution Samples")
plt.show()
```



(b) Use the Acceptance-Rejection Method to obtain the samples from the standard Normal distribution N (0, 1). You are required to show the correctness of your algorithm in theory

[ ]:
```
import numpy as np
import matplotlib.pyplot as plt
```

```python
n_sim = 100000
sample_list = []

# As Normal values from -\infty to +\infty, we use -4 to 4 to cover most of the␣
↪values
value_range = [-4, 4]

def normal_pdf(x, mu=0, sigma=1):
    return np.exp(-0.5 * ((x - mu) / sigma)**2) / (sigma * np.sqrt(2 * np.pi))

def uniform_pdf(x, min, max):
    return 1 / (max - min)

# the maximum value of normal_pdf(x) is normal_pdf(0) while the PDF of Unif is␣
↪1 / (max - min) which is a constant
c = normal_pdf(0) / (1 / (value_range[1] - value_range[0]))

for i in range(n_sim):
    # generate x and y from uniform distribution, the reference distribution is␣
↪Unif
    x = np.random.uniform(value_range[0], value_range[1])
    U = np.random.uniform(0, 1)
    if U < normal_pdf(x) / (c * uniform_pdf(x, value_range[0], value_range[1])):
        sample_list.append(x)

sample_array = np.array(sample_list)

x = np.linspace(value_range[0], value_range[1], 200)
pdf = normal_pdf(x)

plt.hist(sample_array, bins=50, density=True, color="skyblue", alpha=0.7,␣
↪label="Simulated Samples")
plt.plot(x, pdf, color="red", linewidth=2, label="Theoretical PDF")

plt.legend()
plt.xlabel("Value")
plt.ylabel("Density")
plt.title("Histogram of Normal Distribution Samples")
plt.show()
```



(c) Both the Acceptance-Rejection Method and Box-Mulller Method can obtain the samples from the standard Normal distribution N (0,1). Discuss the pros and cons of such two methods

Method	Pros	Cons
Acceptance-Rejection	Can be easily applied to any distribution, as long as the PDF is known	The efficiency is low, as it need to generate large number of samples to get the desired number of samples, so will cost much more time
Box-Muller	The speed is much faster than the Acceptance-Rejection	The inverse step is needed, which is hard to calculate sometimes

(d) Use the importance sampling method to evaluate the probability of rare event c = P(Y > 8), where Y N(0, 1)

[ ]:
```
import numpy as np

samples = 50000
```

```python
indicator = np.zeros(samples)

# generate samples from Uniform distribution
x = np.random.uniform(-10, 10, samples)

# calculate the weights
weights = np.exp(-(x ** 2) / 2) / np.sqrt(2 * np.pi)
indicator[x > 8] = 1

prob = np.sum(weights * indicator) / np.sum(weights)

print("The probability of rare event c = P(Y > 8) is:", prob)
```

```
The probability of rare event c = P(Y > 8) is: 6.175296855207644e-16
```