Applied Stochastic Assignment 1

Manyara Baraka - mbaraka

1 Probability Density Function and Applications

1.1

Normalization

$$\int_0^\infty \lambda^2 y e^{-\lambda y}\, dy = 1$$

Integration by parts formula:

$$\int u dv = uv - \int v du$$

let u = y and dv = $e^{-\lambda y}$
Rewriting the question:

$$\lambda^2 \int_0^\infty y e^{-\lambda y} = 1$$

solve for du by differentiating:

$$du = 1dy$$

solve for v by integrating dv

$$v = \int e^{-\lambda y} = \frac{-e^{-\lambda y}}{\lambda}$$

Substituting to the integration by parts:

$$= y \cdot -\frac{e^{\lambda} y}{\lambda} - \int -\frac{e^{-\lambda y}}{\lambda}$$

$$= y \cdot -\frac{e^{\lambda} y}{\lambda} + \frac{1}{\lambda} \int e^{\lambda - y} dy$$

Integrate the remaining integral

$$= -\frac{y e^{\lambda} y}{\lambda} - \frac{1}{\lambda^2} e^{\lambda - y} dy$$

Multiplying with the λ^2:

$$= \lambda^2 (-\frac{y e^{\lambda} y}{\lambda}) - \lambda^2 (\frac{1}{\lambda^2} e^{\lambda - y}) = -y \lambda e^{\lambda y} - e^{\lambda - y}$$

$$\left[-y\lambda e^{\lambda y} - e^{\lambda - y}\right]_0^\infty = [(0-0)-(0-1)] = 1$$

Therefore, the normalization condition is satisfied.

Concavity and Convexity

To know convexity and Concavity we evaluate the second derivative:

First Derivative of $g(y)$

Given:

$$g(y) = \lambda^2 y e^{-\lambda y}.$$

Use the product rule:

$$g'(y) = \frac{d}{dy}\left(\lambda^2 y e^{-\lambda y}\right).$$

Using the product rule on $ye^{-\lambda y}$:

$$\frac{d}{dy}\left(ye^{-\lambda y}\right) = e^{-\lambda y} + y\frac{d}{dy}\left(e^{-\lambda y}\right).$$

We get:

$$g'(y) = \lambda^2\left(e^{-\lambda y} + y(-\lambda e^{-\lambda y})\right),$$

$$g'(y) = \lambda^2 e^{-\lambda y}(1-\lambda y).$$

Second Derivative of $g(y)$

$$g''(y) = \frac{d}{dy}\left[\lambda^2 e^{-\lambda y}(1-\lambda y)\right].$$

Use product rule:

$$g''(y) = \lambda^2\left[\frac{d}{dy}\left(e^{-\lambda y}\right)(1-\lambda y) + e^{-\lambda y}\frac{d}{dy}(1-\lambda y)\right].$$

Substitute the derivatives:

$$g''(y) = \lambda^2\left[(-\lambda e^{-\lambda y})(1-\lambda y) + e^{-\lambda y}(-\lambda)\right].$$

$$g''(y) = -\lambda^3 e^{-\lambda y}(1-\lambda y) - \lambda^2 e^{-\lambda y}.$$

Simplify the equation

$$g''(y) = -\lambda^2 e^{-\lambda y}(\lambda(1-\lambda y)+1),$$

$$g''(y) = -\lambda^2 e^{-\lambda y}(1-\lambda y+1).$$

$$g''(y) = -\lambda^2 e^{-\lambda y}(2 - \lambda y).$$

Is it Concave or Convex

$$g''(y) = -\lambda^2 e^{-\lambda y}(2 - \lambda y).$$

- $e^{-\lambda y}$ will always positive for all $y \geq 0$.
- $-\lambda^2$ is negative if $\lambda > 0$.
- $(2 - \lambda y)$ changes sign depending on the value of y:
  - If $y < \frac{2}{\lambda}$, then $2 - \lambda y > 0$.
  - If $y > \frac{2}{\lambda}$, then $2 - \lambda y < 0$.

Therefore the Function is:

- **Concave**
  if $y < \frac{2}{\lambda}$, $g''(y)$ is negative.

  This implies that shorter outages occur with high likelihood since the probability decreases.
- **Convex**
  if $y > \frac{2}{\lambda}$, $g''(y)$ is positive.
  This implies that prolonged outages have a rapid decreasing likelihood.

Mode of Distribution

To find the mode of distribution we need to find the critical points by setting the first derivative to zero.

Find the First Derivative From the solution above

$$g'(y) = \lambda^2 e^{-\lambda y}(1 - \lambda y).$$

Find Critical Points
Equate it to zero:

$$g'(y) = \lambda^2 e^{-\lambda y}(1 - \lambda y) = 0.$$

Since $\lambda^2 e^{-\lambda y}$ is never zero for $y \geq 0$, we focus on the term $1 - \lambda y$:

$$1 - \lambda y = 0.$$

Solving for y:

$$y = \frac{1}{\lambda}.$$

Find the Maximum or Minimum

To Know whether the critical point corresponds to a maximum or a minimum, we examine the second derivative:
From the question above:

$$g''(y) = -\lambda^2 e^{-\lambda y}(2 - \lambda y).$$

Substitute $y = \frac{1}{\lambda}$ into the second derivative:

$$g''\left(\frac{1}{\lambda}\right) = -\lambda^2 e^{-1}\left(2 - \lambda\frac{1}{\lambda}\right),$$

$$g''\left(\frac{1}{\lambda}\right) = -\lambda^2 e^{-1}(2 - 1),$$

$$g''\left(\frac{1}{\lambda}\right) = -\lambda^2 e^{-1}.$$

Since λ^2 and e^{-1} are positive, $-\lambda^2 e^{-1}$ is negative. Therefore, $g''\left(\frac{1}{\lambda}\right) < 0$, indicating that the function has a **maximum** at $y = \frac{1}{\lambda}$.

For $\lambda = 0.5$:

$$y = \frac{1}{\lambda} = \frac{1}{0.5} = 2.$$

For $\lambda = 2$:

$$y = \frac{1}{\lambda} = \frac{1}{2} = 0.5.$$

Real-Life Scenarios

For $\lambda = 0.5$ (Mode = 2):

- A smaller λ implies a more spread-out distribution. The mode $y = 2$ suggests that outages are more likely to be prolonged, indicating that longer outages are the most probable duration.

For $\lambda = 2$ (Mode = 0.5):

- A larger λ results in a more concentrated distribution around shorter durations. The mode $y = 0.5$ indicates that the most likely duration of an outage is brief.

1.2

Normalization Condition

Show that:

$$\int_0^\infty \int_0^\infty e^{-(x+y)}\, dx\, dy = 1$$

The integral separates:

$$\int_0^\infty \int_0^\infty e^{-(x+y)}\, dx\, dy = \left(\int_0^\infty e^{-x}\, dx\right)\left(\int_0^\infty e^{-y}\, dy\right)$$

Each integral is 1:

$$\int_0^\infty e^{-x}\, dx = 1, \quad \int_0^\infty e^{-y}\, dy = 1$$

$f(x, y)$ is a valid PDF.

Probability That Both Plants Produce More Than 1 Ton

We need to compute $P(X > 1, Y > 1)$:

$$P(X > 1, Y > 1) = \int_1^\infty \int_1^\infty e^{-(x+y)}\, dx\, dy$$

Separating the integrals:

$$P(X > 1, Y > 1) = \left(\int_1^\infty e^{-x}\, dx\right)\left(\int_1^\infty e^{-y}\, dy\right)$$

Each integral evaluates to e^{-1}, so:

$$P(X > 1, Y > 1) = e^{-1} \times e^{-1} = e^{-2}$$

Probability That Plant A Produces More Than 2 Tons and Plant B Produces Less Than 2 Tons

We need to compute $P(X > 2, Y < 2)$:

$$P(X > 2, Y < 2) = \int_2^\infty \int_0^2 e^{-(x+y)}\, dy\, dx$$

The inner integral is:

$$\int_0^2 e^{-y}\, dy = 1 - e^{-2}$$

The outer integral is:

$$\int_2^\infty e^{-x}\, dx = e^{-2}$$

Thus:

$$P(X > 2, Y < 2) = e^{-2} \times (1 - e^{-2}) = e^{-2} - e^{-4}$$

Production Balance

A higher probability will indicate the tendency of plant A to outperform its target while plant B is underperforming. This information can guide the management in resource management, and risk management and adjust targets or modify the operational strategies to gain balances in the production.

2 EigenValues and Stability analysis

Eigenvalues and Stability

We find the eigenvalues by solving the characteristic equation:

$$\det(B - \lambda I) = 0$$

First, subtract λI from B:

$$B - \lambda I = \begin{pmatrix} 4-\lambda & 2 & 1 \\ 1 & 3-\lambda & 2 \\ 0 & 1 & 2-\lambda \end{pmatrix}$$

Now, calculate the determinant:

$$\det(B - \lambda I) = (4-\lambda) \begin{vmatrix} 3-\lambda & 2 \\ 1 & 2-\lambda \end{vmatrix} - 2 \begin{vmatrix} 1 & 2 \\ 0 & 2-\lambda \end{vmatrix} + 1 \begin{vmatrix} 1 & 3-\lambda \\ 0 & 1 \end{vmatrix}$$

The smaller determinants are:

$$\det \begin{pmatrix} 3-\lambda & 2 \\ 1 & 2-\lambda \end{pmatrix} = \lambda^2 - 5\lambda + 4$$

$$\det \begin{pmatrix} 1 & 2 \\ 0 & 2-\lambda \end{pmatrix} = 2 - \lambda$$

$$\det \begin{pmatrix} 1 & 3-\lambda \\ 0 & 1 \end{pmatrix} = 1$$

Substitute these into the determinant expression:

$$\begin{aligned} \det(B - \lambda I) &= (4-\lambda)(\lambda^2 - 5\lambda + 4) - 2(2-\lambda) + 1 \\ &= -\lambda^3 + 9\lambda^2 - 22\lambda + 13 \end{aligned}$$

The characteristic equation is:

$$-\lambda^3 + 9\lambda^2 - 22\lambda + 13 = 0$$

The eigenvalues (using numerical methods) are approximately:

$$\lambda_1 \approx 0.94, \quad \lambda_2 \approx 2.59, \quad \lambda_3 \approx 5.47$$

Since all eigenvalues have positive real parts, the system is **not stable**.

(b) Impact of Modifying Off-Diagonal Elements

The off-diagonal elements represent the interactions between different features in the system. Increasing these values can push the eigenvalues further into the positive real axis, destabilizing the system. In the context of machine learning, this can result in issues like **exploding gradients**, where the model's parameters grow too large during training, making the system harder to optimize.

3 Markov Chain

State Distribution After One Transition

The initial state distribution:

$$\pi_0 = (0.5, 0.3, 0.2)$$

Transition matrix Q:

$$Q = \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

Therefore:

$$\pi_1 = \pi_0 Q$$

Substitute with the given values

$$\pi_1 = \begin{pmatrix} 0.5 & 0.3 & 0.2 \end{pmatrix} \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix}$$

Matrix multiplication:

$$\pi_1 = \begin{pmatrix} 0.5(0.7) + 0.3(0.3) + 0.2(0.2) & 0.5(0.2) + 0.3(0.4) + 0.2(0.5) & 0.5(0.1) + 0.3(0.3) + 0.2(0.3) \end{pmatrix}$$

$$\pi_1 = \begin{pmatrix} 0.35 + 0.09 + 0.04 & 0.1 + 0.12 + 0.1 & 0.05 + 0.09 + 0.06 \end{pmatrix}$$

After one distribution:

$$\pi_1 = \begin{pmatrix} 0.48 & 0.32 & 0.2 \end{pmatrix}$$

Steady-State Distribution

We solve:

$$\pi Q = \pi$$

Substituting:

$$\begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \end{pmatrix} \begin{pmatrix} 0.7 & 0.2 & 0.1 \\ 0.3 & 0.4 & 0.3 \\ 0.2 & 0.5 & 0.3 \end{pmatrix} = \begin{pmatrix} \pi_1 & \pi_2 & \pi_3 \end{pmatrix}$$

Then we get these equations:

$$\pi_1 = 0.7\pi_1 + 0.3\pi_2 + 0.2\pi_3$$

$$\pi_2 = 0.2\pi_1 + 0.4\pi_2 + 0.5\pi_3$$

$$\pi_3 = 0.1\pi_1 + 0.3\pi_2 + 0.3\pi_3$$

The of the probabilities must equal 1:

$$\pi_1 + \pi_2 + \pi_3 = 1$$

The approximate value of the π:

$$\pi \approx (0.38, 0.35, 0.27)$$

Steady state impact on Employees

The distribution above indicates that the employees are:

- 38% are in High Performance state,
- 35% are in Moderate Performance state
- 27% are in Low Performance state.

The information above shows that most employees are in a high-performance state and moderate performance and the distribution is well distributed for the organization. The information in the organization can be used to focus on necessary training and decisions to improve the employee's performance or improve it. Additionally to know the state of the organization the information is key in making decisions in regards to employee relations.

4 Matrix/Vector Manipulation

Determine the Convexity and Concavity of the function

The eigenvalues of the matrix M are analyzed to determine whether the quadratic form is convex or concave. This is because the Eigenvalues give us information about the Matrix behavior.

For Convexity
The quadratic function $f(x) = x^T M x$ is convex if all **eigenvalues of M are non-negatives**. That is:

All eigenvalues of M must be non-negative (i.e., $\lambda_i \geq 0$ for all i).

For Concavity The quadratic function $f(x) = x^T M x$ is concave if **eigenvalues of M are non-positives**. That is:

All eigenvalues of M must be non-positive (i.e., $\lambda_i \leq 0$ for all i).

Indefinite Curvature: Neither concave nor convex If M has both positive and negative eigenvalues, the function $f(x) = x^T Mx$ is neither convex nor concave.

Connection to Optimization in Machine Learning

In optimization:

- **Convex functions Loss Function**
  Convex Functions are easy to minimize, additionally, the loss function is ideally convex in models like support vector machines (SVMs). This indicates that there is only one minimum point, and with this optimization is made easier. Having a convex form ensures that we can find the global minimum efficiently.
- **Concave Loss functions**
  Convex Functions are easy to maximize since they have a single maximum, additionally they often tend to arise in conditions where we need/want to maximize conditions (like rewarding in reinforcement learning).
- **Non-convex Loss Functions**
  This is usually experienced in deep neural networks where the loss function is mostly non-convex. This indicates the availability of multiple minima that can lead to complexity in optimization.

5 Basic Probabiltity

- **(a) The probability of drawing a heart.**
  A deck has 13 hearts of the 52 cards:
  $$P(Heart) = \frac{13}{52} = \frac{1}{4}$$
- **(b) The probability of drawing a face card (Jack, Queen, King).**
  Each of the 3 face cards are in each suits, therefore:
  $$3 \times 4 = 12\text{Face cards}$$
  $$P(\text{Face card}) = \frac{12}{52} = \frac{3}{13}$$

6 Conditional Probability

Probability that the car customer selected insures one car and its not a sports car.

- Probability of more than one car $P(M) = 0.70$
  The probability of **one car**
  $$P(M^c) = 1 - P(M) = 1 - 0.70 = 0.30$$

- Probability of insuring sports car $P(S) = 0.2$
- Probability of insuring a sports car and they insure more than one car $P(S|M) = 0.15$

  The probability of a sports car given more than one car

$$P(S|M) = P(S|M) \cdot P(M) = 0.15 \times 0.70 = 0.105$$

Probability of insuring a sports car and only one car

$$P(S|M^c) = P(S) - P(S|M) = 0.20 - 0.105 = 0.095$$

Probability of insuring one car and it is a sports car

$$P(M^c|S^c) = P(M^c) - P(S|M^c) = 0.30 - 0.095 = 0.205$$

7 Bayes' Theorem

Probability they have a disease given they tested positive $P(D|P)$

The details given:

- Sensitivity *True Positive* - Positive and you have the disease $P(P|D) = 0.98$
- Specificity *True negative* - $P(\overline{P}|\overline{D}) = 0.97$

$$P(P|\overline{D}) = 1 - 0.97 = 0.03$$

- Having the disease: P(D) = 0.005

$$P(\overline{D}) = 1 - 0.005 = 0.995$$

$$P(D|P) = \frac{P(P|D) \cdot P(D)}{P(P)}$$

$$P(P) = P(P|D) \cdot P(D) + P(P|\overline{D}) \cdot P(\overline{D})$$
$$= (0.98 \times 0.005) + (0.03 \times 0.995) = 0.03475$$

Using Bayes Theorem to solve the probability of having the disease if you tested positive:

$$P(D|P) = \frac{0.98 \times 0.005}{0.03475} = 0.141$$

Probablity of default given that the system predicted to defaulting

Details given:

- $P(PredictDefault|Default) = 0.95$
- $P(PredictDefault|NoDefault) = 0.01$
- $P(Default) = 0.005$
- $P(NoDefault) = 1 - 0.005 = 0.995$

$$P(Default|PredictDefault) = \frac{P(PredictDefault|Default) \cdot P(Default)}{P(PredictDefault)}$$

$$P(PredictDefault) = P(PredictDefault|Default) \cdot P(Default) + P(PredictDefault|NoDefault) \cdot P(NoDefa$$

$$= (0.95 \times 0.005) + (0.01 \times 0.995) = 0.0147$$

Using Bayes theorem to solve:

$$P(Default|PredictDefault) = \frac{0.95 \times 0.005}{0.0147} = 0.323$$

8 Independence of Events, Inclusion-Exclusion Principles, Mutual Exclusivity

Prove for A *and* B being independent so are:

1. **A^c and B**

$$P(A \cap B) = P(A) \cdot P(B)$$
$$P(A^c \cap B) = P(A^c) \cdot P(B)$$
$$P(A^c \cap B) = P(B) - P(A \cap B)$$
$$= P(B) - P(A) \cdot P(B)$$
$$= P(B)(1 - P(A)$$
$$= P(B) \cdot P(A^c)$$

2. **A and B^c**

$$P(A) \cdot P(B^c) = P(A) \cdot P(B^c)$$
$$= P(A) - P(A \cap B)$$
$$= P(A) - P(A) \cdot P(B)$$
$$= P(A)(1 - P(B))$$
$$= P(A) \cdot P(B^c)$$

3. A^c **and** B^c

$$\begin{aligned}P(A^c \cap B^c) &= P(A^c) \cdot P(B^c)\\ &= 1 - P(A \cup B)\\ P(A \cup B) &= P(A) + P(B) - P(A \cap B)\\ &= 1 - [P(A) + P(B) - P(A) \cdot P(B)]\\ &= 1 - P(A) - P(B) + P(A) \cdot P(B)\\ &= (1 - P(A) \cdot (1 - P(B)\\ &= P(A^c) \cdot P(B^c)\end{aligned}$$

Students that take either robotics or ISS

- Robotic students (R) = 15
- ISS students (S) = 10
- Both students $(R \cap S) = 5$

Use the inclusion-exclusion principle:

$$\begin{aligned}P(R \cup S) &= P(R) + P(S) - P(R \cap S)\\ &= 15 + 10 - 5\\ &= 20\end{aligned}$$

Rolling a dice:

- Small prize: Event A Rolling 1 or 2
- Medium prize: Event B Rolling 3,4 or 5
- Large prize: Event C Rolling 6

1. **Events A and B Mutually exclusive**

   Yes, events A and B are mutually exclusive since they cannot co-occur, you cannot roll Event A 1, 2 and Event B 3, 4, 5 on a single dice roll.

2. **Event A and C Mutually exclusice**

   Yes, events A and C are mutually exclusive the y cannot co-occur on one dice roll.

3. **Probability of Event A and B** Since both events are mutually exclusive then:

$$\begin{aligned}P(A \cup B) &= P(A) + P(B)\\ &= \frac{2}{6} + \frac{3}{6}\\ &= \frac{5}{6}\end{aligned}$$

9 Combinatrics

1. **Ways to Arrange ALIGORITHM**
   For the vowels to come together it means will have this new word LGRTHM(AIO) those are 7 units, 6 consonants, and vowels together make it one unit:

   To arrange the 87 units:
   $$7! = 5,040$$
   Arranging the vowels inside the vowel block (A I O):
   $$3! = 6$$
   The complete number of re-arranging the word will be:
   $$8! \times 3! = 5,040 \times 6 = 30,240$$

2. **No of committees and male feuding**
   Males: 7
   Females: 5

   - Committee with 2 females and 3 males from the 5 and 7 respectively.
     Choose 2 females from 5
     $$\binom{5}{2} = \frac{5!}{2!(5-2)!} = 10$$
     Choose 3males from 7
     $$\binom{7}{3} = \frac{7!}{3!(7-3)!} = 35$$
     Complete set of committee
     $$10 \times 35 = 350$$
   - When 2 males Feaud and refuse to serve together
     If 2 males feud it means we will need to choose 1 from the remaining 5:
     $$\binom{5}{1} = \frac{5!}{1!(5-1)!} = 5$$
     Total number of committees with the complete 10 for females:
     $$10 \times 5 = 50$$
     $$350 - 50 = 300$$

3. **Kigali Census**
   Details of the Kigali Census:

   - Total population: 1.75 million
   - 55% population are female: 0.9625 million
   - 60% of the male population is aged 35yrears and above
   - mean μ of hair: 52,500
   - Standard deviation of hair: σ: 10,000

   (a) Expected Number of men aged 35 years and above

$$Number of male = 1.75 \times 0.45 = 0.7875$$

$$= 787,500$$

$$Male aged 35+ = 787,500 \times 0.60$$

$$= 472,500$$

   (b) Range for 68% of hair count for men 35 +
   For normal distribution 68% of data lies between on standard deviation of the mean, that is:

$$Range = [\mu - \sigma, \mu + \sigma]$$

$$Range = [52,500 - 10,000, 52,500 + 10,000]$$

$$Range = [42,500 - 62,500]$$

   (c) Pigeion Hole Principle
   For discrete hair count we use the range:

$$= 62,500 - 42,500 + 1 = 20,001$$

   For Pigeion principle, we work out with:

$$\frac{472,500}{20,001} = 23.62$$

10 Law of Total Probability

The factory produces:

- Type A
  $P(A) = 0.5$, $P(D|A) = 0.01$
- Type B
  $P(B) = 0.3$, $P(D|B) = 0.02$

- Type C
  $P(C) = 0.2$, $P(D|C) = 0.03$

1. **Probability the gadget is Defective**

$$P(D) = P(D|A) \cdot P(A) + P(D|B) \cdot P(B) + P(D|C) \cdot P(C)$$

$$= (0.01 * 0.05) + (0.02 * 0.3) + (0.03 * 0.2)$$

$$= 0.017$$

2. **Defective is from Japan** Samsung factories production:

   - Korea
     $P(K) = 0.5$, $P(D|K) = 0.02$
   - Japan
     $P(J) = 0.3$, $P(D|J) = 0.04$
   - USA
     $P(U) = 0.2$, $P(D|U) = 0.05$

The probability that the defective and its from Japan:

$$P(J|D) = \frac{P(D|J) \cdot P(J)}{P(D)}$$

$$P(D) = P(D|K) \cdot P(K) + P(D|J) \cdot P(J) + P(D|U) \cdot P(U)$$
$$= (0.02 \times 0.5) + (0.04 \times 0.3) + (0.05 \times 0.2)$$
$$= 0.375$$

simulation

September 14, 2024

[10]:
```python
# Import the libraries
import random
import matplotlib.pyplot as plt

#Define the cards and create the tuple to hold the whole deck
suits = ['hearts', 'diamonds', 'clubs', 'spades']
ranks = ['2', '3', '4', '5', '6', '7', '8', '9', '10', 'Jack', 'Queen', 'King',␣
↪'Ace']
deck = [(rank, suit) for suit in suits for rank in ranks]

# What is known
total_cards = 52
hearts_count = 13
face_card_count = 12

heart_prob = hearts_count / total_cards
face_prob = face_card_count / total_cards
```

[13]:
```python
print(f'Probability of picking the heart: \n {heart_prob}')
print(f'Probability of picking face card: \n {face_prob:.2f}')
```

```
Probability of picking the heart:
 0.25
Probability of picking face card:
 0.23
```

[14]:
```python
# Calculate probabilities in a sample
def calculate_sample_probabilities(sample):
    hearts_in_sample = sum(1 for card in sample if card[1] == 'hearts')
    face_cards_in_sample = sum(1 for card in sample if card[0] in ['Jack',␣
↪'Queen', 'King'])
    return hearts_in_sample / len(sample), face_cards_in_sample / len(sample)
```

[17]:
```python
# Simulate drawing cards
sample_size_range = list(range(1, 53))
prob_heart_simulation = []
prob_face_simulation = []
```

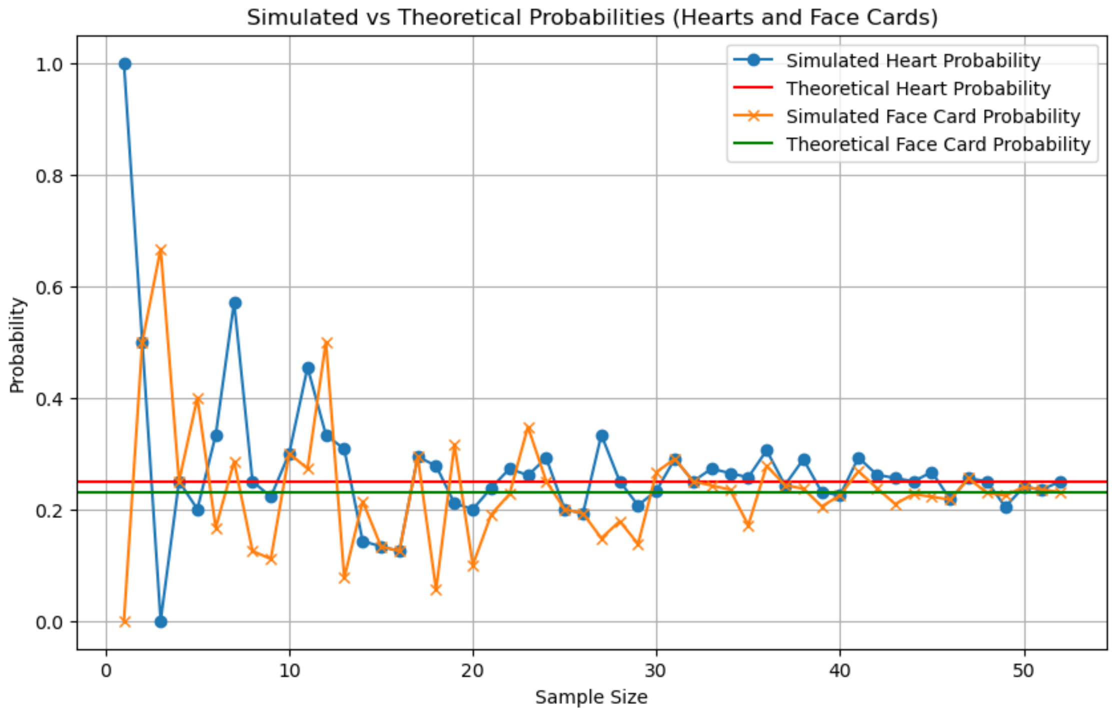
```python
# Run simulations for sample sizes from 1 to 52
for sample_size in sample_size_range:
    sample = random.sample(deck, sample_size)
    prob_heart, prob_face = calculate_sample_probabilities(sample)
    prob_heart_simulation.append(prob_heart)
    prob_face_simulation.append(prob_face)

# Plot calculated and simulated probabilities for hearts and face cards
plt.figure(figsize=(10, 6))

# Plot hearts probabilities
plt.plot(sample_size_range, prob_heart_simulation, label='Simulated Heart Probability', marker='o')
plt.axhline(y=heart_prob, color='r', linestyle='-', label='Theoretical Heart Probability')

# Plot face card probabilities
plt.plot(sample_size_range, prob_face_simulation, label='Simulated Face Card Probability', marker='x')
plt.axhline(y=face_prob, color='g', linestyle='-', label='Theoretical Face Card Probability')

plt.xlabel('Sample Size')
plt.ylabel('Probability')
plt.title('Simulated vs Theoretical Probabilities (Hearts and Face Cards)')
plt.legend()
plt.grid(True)
plt.show()
```


Simulated vs Theoretical Probabilities (Hearts and Face Cards)
Simulated Heart Probability
Theoretical Heart Probability
Simulated Face Card Probability
Theoretical Face Card Probability
Probability
Sample Size
1.0
0.8
0.6
0.4
0.2
0.0
0
10
20
30
40
50

```
[ ]:
```

```
[ ]:
```