1. (15 points) Suppose a box has 8 balls labeled 1, 2, ..., 5. Two balls are drawn from the box without replacement. Let \( x \) be the number of the first ball, and \( y \) the number of the second ball. Answer the following questions:

(a) (5 points) Calculate the mean and variance of \( y \).

(b) (10 points) Calculate the correlation coefficient of \( x \) and \( y \).

2. (10 points) Suppose a box has 5 balls labeled 1, 2, ..., 5. A sequence of balls are independently drawn from the box with replacement. For each of the outcomes, you win $1 dollar if the number of the ball is greater than or equal to 3, and lose $1 if the number is 1 or 2. Suppose you have $10 dollars before the games start. Answer the following questions:

(a) (5 points) What is your expected “wealth” after the 10 games.

(b) (5 points) What is the probability that you will lose half of your money after the 10 games.

3. (25 points) Let \( x \) and \( y \) have the following joint probability density function (pdf):

\[
f(x, y) = \begin{cases} 
1 & -x < y < x \\
0 & \text{otherwise}, 
\end{cases}
\]

where \( 0 < x < 1 \). Answer the following three questions:

(a) (10 points) Suppose the value of \( x \) is known, calculate the conditional mean and variance of \( y \), i.e., \( E(y|x) \) and \( \text{var}(y|x) \).

(b) (10 points) Calculate the unconditional mean and variance of \( y \), i.e., \( E(y) \) and \( \text{var}(y) \).

(c) (5 points) Calculate the probability that \( y \geq \frac{1}{2} \), i.e., \( Pr(y \geq \frac{1}{2}) \).

---

1. Supporters claim that a new windmill can generate an average of at least 800 kilowatts of power per day. Daily power generation for the windmill is assumed to be normally distributed with a standard deviation of 120 kilowatts. A random sample of 109 days is taken to test this claim against the alternative hypothesis that the true mean is less than 800 kilowatts. The claim will be accepted if the sample mean is 776 kilowatts or more and rejected otherwise.

(a) What is the probability \( \alpha \) of a Type I error using the decision rule if the population mean is in fact 800 kilowatts per day? (5%) 

(b) What is the probability \( \beta \) of a Type II error using this decision rule if the population mean is in fact 740 kilowatts per day? (5%):

(c) Suppose that the same decision rule is used, but with a sample of 200 days rather than 109. Would the value of \( \alpha \) be larger than, smaller than, or the same as that found in (a)? (5%)

(d) Suppose that a sample of 100 observations was taken but that the decision rule was changed so that the claim would be accepted if the sample mean was at least 765 kilowatts. Would the value of \( \beta \) be larger than, smaller than, or the same as that found in (b)? (5%)
2. The number of customers arriving at a supermarket checkout counter over a period of 200 minutes was recorded, yielding the results shown in the table. The average number of customers per minute was 2.3. Test the null hypothesis that the population distribution is Poisson. (10%)

Number of Customers in Minutes: 0 1 2 3 4 5 or more
Observed Frequency : 16 50 51 44 28 11

3. Students were classified according to three parental income groups and also according to three possible score ranges in the SAT examination. One student was chosen randomly from each of the nine cross-classifications, and the grade point average of each sample member at the end of the sophomore year was recorded. The results are shown in the accompanying table.

<table>
<thead>
<tr>
<th>SAT Score</th>
<th>Income Group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>High</td>
</tr>
<tr>
<td>Very high</td>
<td>3.7</td>
</tr>
<tr>
<td>High</td>
<td>3.4</td>
</tr>
<tr>
<td>Moderate</td>
<td>2.9</td>
</tr>
</tbody>
</table>

(a) Set out the analysis of variance table. (10%)
(b) For the two-way analysis of variance model with one observation per cell, write the observation from the ith group and jth block as

\[ X_{ij} = \mu + G_i + B_j + \epsilon_{ij} \]

Consider the observation on moderate income group and high SAT score \((X_{12}=2.8).\) Estimate and interpret \(G_2\) and \(B_2.\) (10%)