Hypothesis testing

Exercise 1 (Balls and bins)**:**
A bin whose content is hidden is supposed to contain 2 blue balls and 2 red balls. You want to show that there is in fact only 1 blue ball and 3 red balls. You propose to test this hypothesis by randomly and simultaneously selecting 2 balls from the bin and decide that you are right if none of them is blue.

1. What is the null hypothesis H_0? the alternative hypothesis H_1?
2. Give the type I and type II error rates.

Exercise 2 (Playing cards)**:**
A deck of playing cards is supposed to be a standard 52-card deck. You want to show that it is actually a 32-card deck (without 2, 3, 4, 5, 6). You propose to test this hypothesis by randomly drawing 4 cards from the deck and observe the presence or absence of 2, 3, 4, 5, 6.

1. What is the null hypothesis H_0? the alternative hypothesis H_1?
2. Give the type I and type II error rates.

Exercise 3 (Exponential model)**:**
The lifetime of a component is supposed to be exponential with parameter $\theta > 0$. You observe n i.i.d. samples.

1. Propose a uniformly most powerful (UMP) test of the null hypothesis $H_0 = \{\theta = 1\}$ against the alternative hypothesis $H_1 = \{\theta = 10\}$ at level α.
2. Specify the test for $n = 1$ and $\alpha = 10\%$.
   What is the power of this test?
3. Propose a UMP test of the null hypothesis $H_0 = \{\theta \le 1\}$ against the alternative hypothesis $H_1 = \{\theta > 1\}$ at level α, and compare with the test of question 1.

Exercise 4 (Beta model)**:**
Some observe n i.i.d. samples of some physical quantity, that have a Beta distribution with parameters (θ, θ) for some $\theta > 0$, that is

$$p_\theta(x) \propto \prod_{i=1}^n x_i^{\theta-1}(1-x_i)^{\theta-1}.$$

1. Propose a uniformly most powerful (UMP) test of the null hypothesis $H_0 = \{\theta \le 1\}$ against the alternative hypothesis $H_1 = \{\theta > 1\}$ at level α.
2. Specify the test for $n = 1$ and $\alpha = 10\%$.
   What is the power of this test for $\theta = 2$?

Exercise 5 (Speed limitation):
The average number of car accidents per year is equal to 1200 in some country. It is decided to reduce the maximum speed. You must decide whether the new speed limitation is efficient. You assume that the numbers of car accidents per month are i.i.d. Poisson with parameter θ and would like to test two hypotheses, $A = \{\theta = 100\}$ and $B = \{\theta = 90\}$.

1. Choose the relevant null hypothesis if you represent (a) the victims of car accidents, and (b) the drivers.
2. Using a Gaussian approximation, propose in each case a test at level $\alpha = 1\%$ based on the total number of car accidents after n months.
3. Give the power of the test in each case.
4. The total number of car accidents after one year of speed limitation is equal to 1100. What is your conclusion? Give the p-value associated with each case.

Exercise 6 (Bernoulli model – simple hypotheses):
Let θ be the fraction of electric cars in Paris. You want to test $H_0 = \{\theta_0\}$ with $\theta_0 = \frac{1}{4}$ against $H_1 = \{\theta_1\}$ with $\theta_1 = \frac{1}{3}$ using n i.i.d. observations.

1. Give a uniformly most powerful test at level α.
2. Using a Gaussian approximation, specify the rejection region of the null hypothesis for $\alpha = 5\%$ and $n = 100$ cars. Give the power of the test.
3. Give the result of the test if you get 30 electric cars among $n = 100$.

Exercise 7 (Gaussian model – unknown mean, simple hypotheses):
The daily power consumptions of a company are supposed to be i.i.d. gaussian with unknown mean θ and known variance σ^2. You want to test $H_0 = \{\theta_0\}$ against $H_1 = \{\theta_1\}$ with $\theta_0 > \theta_1$.

1. Give a uniformly most powerful test at level α based on n observations.
2. Give the result of the test for $\theta_0 = 200$, $\theta_1 = 160$, $\sigma = 50$, $\alpha = 1\%$ and an empirical mean consumption of 165 based on $n = 10$ observations.

Exercise 8 (Gaussian model – unknown variance, simple hypotheses):
The daily power consumptions of a company are supposed to be i.i.d. gaussian with known mean μ and unknown variance θ. You want to test $H_0 = \{\theta_0\}$ against $H_1 = \{\theta_1\}$ with $\theta_0 < \theta_1$.

1. Give a uniformly most powerful test at level α based on n observations.
2. Give the result of the test for $\mu = 200$, $\theta_0 = 400$, $\theta_1 = 900$, $\alpha = 1\%$ and the following observations of the daily consumption: $180, 240, 190, 250, 170, 160$.

Exercise 9 (Pareto model – one-tailed test)**:**
The delay X to receive a shipment is supposed to be Pareto with parameter $\theta > 0$:

$$P_\theta(X > x) = \frac{1}{x^\theta} \quad \forall x \geq 1.$$

You want to test $H_0 = \{\theta = 1\}$ against $H_1 = \{\theta > 1\}$.

1. Give a uniformly most powerful test at level α based on n observations.
2. Give the result of the test for $\alpha = 10\%$ and the following observations: $1, 1, 2, 2, 5, 2$.

Exercise 10 (Bernoulli model – two-tailed test)**:**
You observe the fraction θ of 6 obtained with a dice. You want to test $H_0 = \{\theta = \frac{1}{6}\}$ against $H_1 = \{\theta \neq \frac{1}{6}\}$ using n i.i.d. observations.

1. Give a uniformly most powerful test at level α among all unbiased tests[1].
2. Using a Gaussian approximation, specify the rejection region of H_0 for $\alpha = 1\%$ and $n = 100$.
3. Give the result of the test if you get 10 times a 6 over $n = 100$ samples.

Exercise 11 (Lotteries)**:**
Let θ_1 and θ_2 be the winning probabilities of 2 lotteries. You want to test $H_0 = \{\theta_1 = \theta_2\}$ against $H_1 = \{\theta_1 \neq \theta_2\}$ using n_1 i.i.d. observations of the first lottery and n_2 i.i.d. observations of the second lottery.

1. Propose a test based on the respective proportions of wins p_1 and p_2.
2. Using a Gaussian approximation, specify the rejection region of H_0 for $\alpha = 5\%$, $n_1 = 100$ and $n_2 = 50$.
3. Give the result of the test for $p_1 = 0.1$ and $p_2 = 0.2$.

[1] A test of level α is said to be unbiased if the probability of rejecting H_0 is at least α under H_1.