# Hypothesis testing

**Exercise 1** (Balls and bins)**:**

1. $H_0 = \{2 \text{ blue}, 2 \text{ red}\}$, alternative hypothesis $H_1 = \{1 \text{ blue}, 3 \text{ red}\}$.
2. Let $x$ be the number of blue balls.
   Decision:
   $$\delta(x) = 1_{\{x=0\}}$$
   Type I error rate:
   $$\alpha = P_0(X = 0) = \frac{1}{\binom{4}{2}} = \frac{1}{6}$$
   Type II error rate:
   $$\beta = P_1(X > 0) = 1 - P_1(X = 0) = 1 - \frac{\binom{3}{2}}{\binom{4}{2}} = \frac{1}{2}$$

**Exercise 2** (Playing cards)**:**

1. $H_0 = \{52 \text{ cards}\}$, alternative hypothesis $H_1 = \{32 \text{ cards}\}$.
2. Let $x$ be the number of cards in the set $\{2, 3, 4, 5, 6\}$.
   Decision:
   $$\delta(x) = 1_{\{x=0\}}$$
   Type I error rate:
   $$\alpha = P_0(X = 0) = \frac{\binom{32}{4}}{\binom{52}{4}} \approx 0.16.$$
   Type II error rate:
   $$\beta = P_1(X > 0) = 0.$$

**Exercise 3** (Exponential model)**:**
Statistical model:
$$p_\theta(x) = \prod_{i=1}^{n} \theta e^{-\theta x_i}, \quad x_1, \dots, x_n > 0.$$

1. Simple hypothesis, with $\theta_0 = 1$ and $\theta_1 = 2$. Likelihood ratio:
   $$\frac{p_1(x)}{p_0(x)} = e^{-(\theta_1 - \theta_0)S} = e^{-S}, \quad S = \sum_{i=1}^{n} x_i.$$
   Thus test $\delta(x) = 1_{\{S<c\}}$. Level $\alpha = P_0(S < c)$. $S$ has a Gamma distribution with parameters $(\theta, n)$. We get $c = Q(\alpha)$ with $Q$ the quantile function of a Gamma distribution with parameters $(1, n)$.

2. For $n = 1$, $S$ is exponential with parameter $\theta$ and $\alpha = P_0(S < c) = 1 - e^{-c}$ so that $c = -\ln(1-\alpha) \approx 0.1$.
   Type II error rate $\beta = P_{10}(S > c) = e - 10c = (1-\alpha)^{10}$, power $1 - \beta =\approx 65\%$.

3. For each $\theta' > \theta$,
$$\frac{p_{\theta'}(x)}{p_\theta(x)} = e^{-(\theta'-\theta)S}.$$
   Monotonic likelihood ratio in $S$. Thus $\delta$ is also the uniform most powerful test of level $\alpha$ for $H_0 = \{\theta \leq 1\}$ against $H_1 = \{\theta > 1\}$.

**Exercise 4** (Beta model)**:**
Some observe $n$ i.i.d. samples of some physical quantity, that have a Beta distribution with parameters $(\theta, \theta)$ for some $\theta > 0$, that is
$$p_\theta(x) \propto \prod_{i=1}^n x_i^{\theta-1}(1-x_i)^{\theta-1}.$$

1. Likelihood ratio:
$$\frac{p_{\theta'}(x)}{p_\theta(x)} = P^{\theta'-\theta}, \quad P = \prod_{i=1}^n x_i(1-x_i).$$
   Test UMP:
$$\delta(x) = 1_{P>c}.$$
   Level $\alpha = P_1(P > c)$.

2. For $n = 1$, the test takes the form $\delta(x) = 1_{|x-\frac{1}{2}|<c'}$.
   Level $\alpha = P_1(|X - \frac{1}{2}| < c')$. For $\alpha = 10\%$, $c' = \frac{\alpha}{2} = 5\%$.
   Power $1 - \beta(2) = P_2(|X - \frac{1}{2}| < \frac{\alpha}{2})$.
   We get: $1 - \beta(2) = \int_{\frac{1-\alpha}{2}}^{\frac{1+\alpha}{2}} 6x(1-x)\mathrm{d}x = 6\int_{-\frac{\alpha}{2}}^{\frac{\alpha}{2}}(\frac{1}{4} - u^2)\mathrm{d}u = \frac{3}{2}\alpha(1 - \frac{\alpha}{3})$.
   For $\alpha = 10\%$, the power is $\approx 14\%$.

**Exercise 5** (Speed limitation)**:**

1. The focus must be on the hypothesis you want to prove. If this is taken as the alternative hypothesis, a type I error corresponds to the case where you falsely decide that your hypothesis is correct. (a) If you represent the victims of car accidents, you want to prove that speed limitation reduces the number of accidents. So the null hypothesis is that speed limitation has no effect, $H_0 = A$, and the alterative hypothesis is that speed limitation is efficient, $H_1 = B$. A type I error rate occurs if you decide that speed limitation is efficient whereas it is not. (b) If you represent the drivers, you want to prove that speed limitation does *not* reduce the number of accidents. So the null hypothesis is that speed limitation reduces the number of accidents, as many believe, $H_0 = B$, and the alterative hypothesis is that it is not efficient, $H_1 = A$. A type I error rate occurs if you decide that speed limitation not is efficient whereas it is.

2. Let $x = (x_1, \dots, x_n)$ be the numbers of car accidents per month and $S$ their sum. We have:
$$p_\theta(x) = \prod_{i=1}^{n} \frac{\theta^{x_i}}{x_i!} e^{-\theta}$$
so the likelihood ratio is:
$$\frac{p_1(x)}{p_0(x)} \propto \left(\frac{\theta_1}{\theta_0}\right)^S$$
and a UMP test has the form:
$$\delta(x) = 1_{S>c} \quad \text{or} \quad \delta(x) = 1_{S<c},$$
if $\theta_1 > \theta_0$ or $\theta_1 < \theta_0$, respectively.

Using a Gaussian approximation, we have $S \sim \mathcal{N}(n\theta, n\theta)$. In case (a), you want to prove that speed limitation is efficient, so that the test is of the form $S < c$ for some constant $c$. The level of the test is:
$$\alpha = P_A(S < c).$$
Now $S \sim \sqrt{n\theta} Z + n\theta$ with $Z \sim \mathcal{N}(0,1)$, so that:
$$\alpha = P_A(Z < \frac{c - n\theta}{\sqrt{n\theta}}) = P_A(Z > \frac{n\theta - c}{\sqrt{n\theta}})$$
For $Q$ the quantile function of $Z$, we get:
$$\frac{n\theta - c}{\sqrt{n\theta}} = Q(1 - \alpha)$$
that is:
$$c = n\theta - \sqrt{n\theta} Q(1 - \alpha).$$
For $\alpha = 1\%$, we have $Q(1 - \alpha) \approx 2.3$ so that for $n = 12$ and $\theta = 100$ (null hypothesis):
$$c \approx 1120.$$

For the other case,
$$\alpha = P_B(S > c)$$
and
$$c = n\theta + \sqrt{n\theta} Q(1 - \alpha).$$
For $\alpha = 1\%$, $n = 12$ and $\theta = 90$ (null hypothesis):
$$c \approx 1156.$$

3. For case (a),
$$\beta = P_B(S \geq c) = P_B(\sqrt{n\theta} Z + n\theta \geq c) = P_B(Z > \frac{c - n\theta}{\sqrt{n\theta}}).$$

For $c \approx 1120$, $n = 12$ and $\theta = 90$ (alternative hypothesis), we get

$$\beta \approx 0.12.$$

So the power is $1 - \beta \approx 88\%$.

For case (b),

$$\beta = P_A(S \leq c) = P_A(\sqrt{n\theta}Z + n\theta \leq c) = P_A(Z \leq \frac{c - n\theta}{\sqrt{n\theta}}) = P_A(Z > \frac{n\theta - c}{\sqrt{n\theta}})$$

For $c \approx 1156$, $n = 12$ and $\theta = 100$ (alternative hypothesis), we get

$$\beta \approx 0.10.$$

So the power is $1 - \beta \approx 90\%$.

4. In case (a), you reject the null hypothesis and conclude that the speed limitation is efficient. In case (b), you cannot reject the null hypothesis and thus cannot prove that the speed limitation is inefficient.

   In case (a), the $p$-value is $p = P_A(S < s)$ with $s = 1100$, that is:

$$P_A(Z > \frac{n\theta - s}{\sqrt{n\theta}}).$$

   We get $p = 0.003$.

   In case (b), the $p$-value is $p = P_B(S > s)$ with $s = 1100$, that is:

$$P_B(Z > \frac{s - n\theta}{\sqrt{n\theta}}).$$

   We get $p = 0.27$.

**Exercise 6** (Bernoulli model)**:**

1. Likelihood ratio:

$$\frac{p_1(x)}{p_0(x)} = \frac{\theta_1^S(1-\theta_1)^{n-S}}{\theta_0^S(1-\theta_0)^{n-S}} \propto \left(\frac{\theta_1}{1-\theta_1} \times \frac{1-\theta_0}{\theta_0}\right)^S,$$

   with $S = \sum_{i=1}^n x_i$. Since $\theta_0 < \theta_1$, a uniformly most powerful test is given by:

$$\delta(x) = 1_{S>c},$$

   for some constant $c > 0$.

2. Using a Gaussian approximation, $S \sim \mathcal{N}(n\theta, n\theta(1-\theta))$ so that $S \sim \sqrt{n\theta(1-\theta)}Z + n\theta$ and

$$\alpha = P_0(S > c) = P_0(\sqrt{n\theta(1-\theta)}Z + n\theta > c) = P_0(Z > \frac{c - n\theta}{\sqrt{n\theta(1-\theta)}}).$$

We get:

$$\frac{c - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}} = Q(1-\alpha),$$

that is

$$c = \sqrt{n\theta_0(1-\theta_0)}Q(1-\alpha) + n\theta_0.$$

For $\alpha = 5\%$, $Q(1-\alpha) \approx 1.65$ and $c \approx 32$.

3. We get:

$$\beta = P_1(S \le c) = P_1(Z < \frac{c - n\theta}{\sqrt{n\theta(1-\theta)}}) = P_1(Z > \frac{n\theta - c}{\sqrt{n\theta(1-\theta)}}) = 1 - F(\frac{n\theta_1 - c}{\sqrt{n\theta_1(1-\theta_1)}}).$$

where $F$ is the cumulative distribution function of the standard normal distribution. For $c \approx 32$, we get $\beta \approx 0.42$.

4. For $s = 30$, the null hypothesis cannot be rejected. The $p$-value is:

$$p = P_0(S > s) = 1 - F(\frac{s - n\theta_0}{\sqrt{n\theta_0(1-\theta_0)}}) = 0.13.$$

**Exercise 7** (Gaussian model – unknown mean, simple hypotheses)**:**

1. Likelihood ratio:

$$\frac{p_1(x)}{p_0(x)} \propto \frac{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta_1)^2}}{e^{-\frac{1}{2\sigma^2}\sum_{i=1}^n (x_i-\theta_0)^2}} \propto e^{\frac{1}{\sigma^2}(\theta_1-\theta_0)S}$$

with $S = \sum_{i=1}^n x_i$. Since $\theta_1 < \theta_0$, a uniformly most powerful test is given by:

$$\delta(x) = 1_{S<c},$$

for some constant $c > 0$. Since $S \sim \mathcal{N}(n\theta, n\sigma^2)$, we have $S \sim \sqrt{n\sigma^2}Z + n\theta$ with $Z \sim \mathcal{N}(0,1)$ and

$$\alpha = P_0(S < c) = P_0(\sqrt{n\sigma^2}Z + n\theta < c) = P_0(Z < \frac{c - n\theta}{\sqrt{n\sigma^2}}) = P_0(Z > \frac{n\theta - c}{\sqrt{n\sigma^2}}).$$

We get:

$$\frac{n\theta_0 - c}{\sqrt{n\sigma^2}} = Q(1-\alpha),$$

that is

$$c = n\theta_0 - \sqrt{n\sigma^2}Q(1-\alpha).$$

2. For $\theta_0 = 200$, $\sigma = 50$, $\alpha = 1\%$, we get $c \approx 1636$ so you cannot reject $H_0$. Observe that the result does not depend on $\theta_1$, which only impacts the power of the test.

**Exercise 8** (Gaussian model – unknown variance, simple hypotheses)**:**

1. Likelihood ratio:
$$\frac{p_1(x)}{p_0(x)} \propto \frac{e^{-\frac{1}{2\theta_1}\sum_{i=1}^n (x_i-\mu)^2}}{e^{-\frac{1}{2\theta_0}\sum_{i=1}^n (x_i-\mu)^2}} \propto e^{\frac{1}{2}\left(\frac{1}{\theta_0}-\frac{1}{\theta_1}\right)S}$$
with $S = \sum_{i=1}^n (x_i - \mu)^2$. Since $\theta_1 > \theta_0$, a uniformly most powerful test is given by:
$$\delta(x) = 1_{S>c},$$
for some constant $c > 0$. We have $S \sim \theta Z$ with $Z \sim \chi^2(n)$ and
$$\alpha = P_0(S > c) = P_0(\theta Z > c) = P_0(Z > \frac{c}{\theta}).$$
We get using the quantile function $Q$ of the $\chi^2(n)$ distribution:
$$\frac{c}{\theta_0} = Q(1-\alpha),$$
that is
$$c = \theta_0 Q(1-\alpha).$$
2. For $n = 6$ and $\alpha = 1\%$, we get $Q(1-\alpha) \approx 16.8$ and $c \approx 6720$. Since $S = 7100$, you reject $H_0$.

**Exercise 9** (Pareto model – one-tailed hypothesis)**:**

1. Likelihood for one sample:
$$p_\theta(x) = \frac{\theta}{x^{\theta+1}}, \quad x > 1.$$
Likelihood ratio for any $\theta_1 > \theta_0$:
$$\frac{p_1(x)}{p_0(x)} = \frac{\theta_1}{\theta_0} T(x)^{\theta_0-\theta_1},$$
with $T(x) = \prod_{i=1}^n x_i$. A uniformly most powerful test is given by:
$$\delta(x) = 1_{T(x)<c},$$
for some constant $c > 0$.

Test at level $\alpha$:
$$\alpha = P_0(T(x) < c).$$
Now:
$$\log T(x) = \sum_{i=1}^n \log x_i$$
and for all $t \geq 0$:
$$P(\log X > t) = P(X > e^t) = e^{-t\theta}$$
so that $\log X \sim \mathcal{E}(\theta)$ and $\log T(x) \sim \Gamma(n, \theta)$, that is $\log T(x) \sim \frac{Z}{\theta}$ with $Z \sim \Gamma(n, 1)$. We obtain:
$$\alpha = P_0(Z < \theta \log c),$$

that is, using the quantile function $Q$ of the $\Gamma(n,1)$ distribution:

$$\log c = Q(\alpha),$$

and

$$c = e^{Q(\alpha)}.$$

2. For $n = 6$ and $\alpha = 10\%$, we get $Q(\alpha) \approx 3.15$ and $c \approx 23.4$. You get $T(x) = 30$ so you don't reject $H_0$.

**Exercise 10** (Bernoulli model – two-tailed hypothesis)**:**

1. The model is in the exponential family:

$$p_\theta(x) \propto e^{S\log(\frac{\theta}{1-\theta})}$$

with $S = \sum_{i=1}^n x_i$. This is monotonic so the test has the form:

$$\delta(x) = 1_{S<c_1} + 1_{S>c_2},$$

for some constants $c_1, c_2$ with $c_1 < c_2$. These constants satisfy:

$$\alpha = P_0(S < c_1) + P_0(S > c_2).$$

2. Using a Gaussian approximation, we have $S \sim \mathcal{N}(n\theta, n\theta(1-\theta))$, that is $S \sim \sqrt{n\theta(1-\theta)}Z + n\theta$ with $Z \sim \mathcal{N}(0,1)$. By symmetry, $c_1 = n\theta - \delta$ and $c_2 = n\theta + \delta$ and

$$\alpha = 2P_0(S > n\theta + \delta) = 2P_0(Z > \frac{\delta}{\sqrt{n\theta(1-\theta)}}).$$

Thus, with $\theta_0 = \frac{1}{6}$,

$$\frac{\delta}{\sqrt{n\theta_0(1-\theta_0)}} = Q(1-\frac{\alpha}{2}),$$

that is:

$$\delta = \sqrt{n\theta_0(1-\theta_0)}Q(1-\frac{\alpha}{2}).$$

For $\alpha = 1\%$ and $n = 100$, we get $Q(1-\frac{\alpha}{2}) \approx 2.58$ and $\delta \approx 9.6$ so the reject region is $\{S \le 7\} \cup \{S \ge 27\}$.

3. For $S = 18$, you don't reject $H_0$.

**Exercise 11** (Lotteries)**:**

1. Test:

$$\delta(x) = 1_{\{|p_1-p_2|>c\}}.$$

Level $\alpha$ with $\theta = (\theta_1, \theta_2)$:

$$\alpha = \sup_{\theta:\theta_1=\theta_2} P_\theta(|p_1 - p_2| > c).$$

2. Using a Gaussian approximation, $p_1 \sim \mathcal{N}(\theta_1, \frac{\theta_1(1-\theta_1)}{n_1})$, $p_2 \sim \mathcal{N}(\theta_2, \frac{\theta_2(1-\theta_2)}{n_2})$ and
$$p_1 - p_2 \sim \mathcal{N}(\theta_1 - \theta_2, \frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2}).$$
Thus
$$\alpha = \sup_{\theta:\theta_1=\theta_2} P_\theta(\sqrt{\frac{\theta_1(1-\theta_1)}{n_1} + \frac{\theta_2(1-\theta_2)}{n_2}}|Z| > c) = 2P(\frac{1}{2}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}Z > c).$$
We get:
$$c = Q(1 - \frac{\alpha}{2})\frac{1}{2}\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}.$$
Reject region: $|p_1 - p_2| > c$.
3. We get $Q(1 - \frac{\alpha}{2}) \approx 1.96$ so $c \approx 0.17$. With $p_1 = 0.1$ and $p_2 = 0.2$, you don't reject $H_0$.