Tutorial 11

1. We have a single observation on a random variable X from a distribution with density

$$p(x;\theta) = \begin{cases} e^{-(x-\theta)} & x \geq \theta \\ 0 & x < \theta \end{cases}$$

where $\theta > 0$ is unknown. We test $H_0 : \theta = 0$ against the alternative $H_1 : \theta > 0$ and we reject the null hypothesis if the observed value $x \in [c, +\infty) = \mathcal{R}_{\text{crit}}$ for an appropriate $c > 0$.

   (a) Compute c if the test has significance level $\alpha = 0.05$.

   (b) Determine whether or not this test is uniformly most powerful.

2. Let $X_1, \ldots, X_n$ be i.i.d. with distribution $F(x)$ where

$$F(x) = \begin{cases} 1 - e^{-x^\theta} & x \geq 0 \\ 0 & x < 0 \end{cases} \qquad \theta > 0.$$

Find the most powerful test for $H_0 : \theta = 1$ versus $H_1 : \theta = \theta_1$ for a particular $\theta_1 > 1$. For $\alpha = 0.05$, show that this does not give a UMP test for $H_0 : \theta = 1$ versus $H_1 : \theta > 1$.

3. Let

$$X_i = \begin{pmatrix} X_{i1} \\ X_{i2} \end{pmatrix} \sim N\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix} \right) \qquad i = 1, \ldots, n.$$

and suppose that $X_1, \ldots, X_n$ are independent. Consider the hypothesis test: $H_0 : \mu_1 = \mu_2$ and $\sigma_1 = \sigma_2$ versus the alternative $H_1 : \mu_1 \neq \mu_2$ or $\sigma_1 \neq \sigma_2$. Compute the likelihood ratio test statistic.

4. The $F_{n,m}$ distribution is defined as follows: if $V \sim \chi^2_m$, $W \sim \chi^2_n$ and $V \perp W$, then $F := \frac{W/n}{V/m}$ has $F_{n,m}$ distribution. Let $X_1, \ldots, X_{n_1}$ and $Y_1, \ldots, Y_{n_2}$ be independent exponential $\text{Exp}(\theta)$ and $\text{Exp}(\lambda)$ samples respectively and let $\Delta = \frac{\theta}{\lambda}$.

   (a) Let $f(\alpha)$ denote the value such that $\mathbb{P}(F > f(\alpha)) = \alpha$ where $F \sim F_{2n_1, 2n_2}$. Show that $\left[\frac{\overline{Y}}{\overline{X}} f\left(1 - \frac{\alpha}{2}\right), \frac{\overline{Y}}{\overline{X}} f\left(\frac{\alpha}{2}\right)\right]$ is a confidence interval for Δ with confidence coefficient $1 - \alpha$.

   (b) Show that the test with acceptance region (the region where H_0 is not rejected) given by $[f(1 - \alpha/2), f(\alpha/2)]$ for the test $H_0 : \Delta = 1$ versus $H_1 : \Delta \neq 1$ using test statistic $\widehat{\Delta} = \frac{\overline{X}}{\overline{Y}}$ has size α.

5. Let $X_1, \ldots, X_n$ denote the times (in days) to failure of n similar pieces of equipment which is considered to be an $\text{Exp}(\lambda)$ random sample. Consider the hypothesis $H_0 : \frac{1}{\lambda} = \mu \leq \mu_0$ (the average lifetime is no greater than μ_0).

   (a) Show that the test with critical region $\overline{X} \in \left[\mu_0 \frac{k_{2n,\alpha}}{2n}, +\infty\right)$ where $k_{m,\alpha}$ is the value such that $\mathbb{P}(W > k_{m,\alpha}) = \alpha$ for $W \sim \chi^2_m$, is a size α test.

(b) Give an expression for the power function in terms of the χ^2_{2n} distribution.

(c) Use the central limit theorem to show that $\Phi\left(-\frac{\mu_0 z_\alpha}{\mu} + \frac{\sqrt{n}(\mu-\mu_0)}{\mu}\right)$ is an approximation to the power function of the test in part (a). Here z_α is the value such that $\mathbb{P}(Z > z_\alpha) = \alpha$ for $Z \sim N(0,1)$ and $\Phi(z) = \mathbb{P}(Z \leq x)$.

6. Let $X_1, \ldots, X_n$ be a random sample from Poiss(θ), where θ is unknown.

(a) Construct a UMP level α test for $H_0 : \theta \leq \theta_0$ versus $H_1 : \theta > \theta_0$.

(b) Show that the power function of the test is increasing in θ.

(c) What distribution tables would you need to calculate the power function of the UMP test?

(d) Give an approximate expression, derived using the central limit theorem, for the critical value (above which you reject H_0) if n is large and θ not too close to 0 or $+\infty$.

7. (a) Given a random sample $X_1, \ldots, X_n$ from a distribution with c.d.f. F, let

$$\widehat{F}_n(x) = \frac{1}{n}\sum_{j=1}^{n} \mathbf{1}_{(-\infty,x]}(X_j)$$

denote the empirical distribution. Consider the test of $H_0 : F = F_0$ versus the alternative $H_1 : F \neq F_0$ for a given F_0. Let $D_n = \sup_{x\in\mathbb{R}} |\widehat{F}_n(x) - F_0(x)|$ and consider the test: reject H_0 if and only if $D_n \geq k_\alpha$ for k_α such that $\mathbb{P}_{F_0}(D_n \geq k_\alpha) = \alpha$ under H_0. (Recall that asymptotically D_n has the Kolmogorov distribution).

Show that the *power* of this test (for a true distribution F), $\beta(F)$, is bounded below by

$$\beta(F) \geq \sup_x \mathbb{P}_F\left(|\widehat{F}_n(x) - F_0(x)| \geq k_\alpha\right).$$

(b) For $n = 80$, obtain an approximation $\widetilde{k_{0.1}}$ for $k_{0.1}$ and see how well it approximates the true value:

$$k_{0.1} \simeq \frac{1.2}{\sqrt{n}} \simeq 0.134$$

(c) Again, take $\alpha = 0.10$ and $n = 80$. Let F_0 be the $N(0,1)$ c.d.f. and

$$F(x) = \frac{1}{1+\exp\left\{-\frac{x}{\tau}\right\}} \qquad -\infty < x < +\infty \qquad \tau = \frac{\sqrt{3}}{\pi}.$$

With this choice of τ, this is the *logistic* distribution with mean zero and variance 1. Evaluate the lower bound $\mathbb{P}_F(|\widehat{F}_n(x) - F_0(x)| \geq k_\alpha)$ for $\alpha = 0.10$, $n = 80$ and $x = 1.5$ using the normal approximation to the binomial distribution of $n\widehat{F}(x)$ and the approximate critical value of the previous part. (the value of 1.2 may be obtained from tables of the Kolmogorov Smirnov distribution). If F_0 is the c.d.f. for $N(0,1)$ then $F_0(1.5) \simeq 0.93$.

(d) Show that if $F \neq F_0$ and F and F_0 are continuous, then the power of this test tends to 1 as $n \to +\infty$. You may use the fact that $\sqrt{n}D_n$ under the null hypothesis converges to the Kolmogorov Smirnov distribution. In particular, $\sqrt{n}k_{0.1} \overset{n\to+\infty}{\longrightarrow} c_{0.1} = 1.2$.

8. Again, let $X_1, \ldots, X_n$ be a random sample from a distribution with continuous c.d.f. F and let $\widehat{F}_n$ denote the empirical distribution. Let $\psi : (0,1) \to (0, +\infty)$ and $\alpha > 0$. Define the statistics:

$$S_{\psi,\alpha} = \sup_x \psi(F_0(x))|\widehat{F}(x) - F_0(x)|^\alpha$$

$$T_{\psi,\alpha} = \sup_x \psi(\widehat{F}_n(x))|\widehat{F}(x) - F_0(x)|^\alpha$$

$$U_{\psi,\alpha} = \int \psi(F_0(x))|\widehat{F}_n(x) - F_0(x)|^\alpha F_0(dx)$$

$$V_{\psi,\alpha} = \int \psi(\widehat{F}_n(x))|\widehat{F}_n(x) - F_0(x)|^\alpha \widehat{F}_n(dx)$$

For each of these statistics show that the distribution under $H_0 : F = F_0$, does not depend on F_0 (continuous).