Kyle McCorav

ACM 157 Set 5

1. $X_1, \dots, X_n \sim U[0,\theta]$ $H_0: \theta = \frac{1}{2}$ $H_1: \theta > \frac{1}{2}$ $R = \{X : X_{(n)} > c\}$

a)

$P(\text{Type 1 error}) = P(X \in R \mid \theta \in \Theta_0) = P(X_{(n)} > c \mid \theta = \frac{1}{2})$
$P(\text{Type 2 error}) = P(X \notin R \mid \theta \in \Theta_1) = 1 - P(X_{(n)} > c \mid \theta > \frac{1}{2})$

$\beta(\theta) = P(X \in R \mid \theta) = P(X_{(n)} > c \mid \theta) = 1 - \left(\frac{c}{\theta}\right)^n$ if $\theta \in [0,\theta]$
(1 if $c < 0$, 0 if $c > \theta$)

b)

$\alpha = \sup_{\theta \in \Theta_0} \beta(\theta) = \sup_{\theta = 1/2} 1 - \left(\frac{c}{\theta}\right)^n = 1 - (2c)^n$

$c = \frac{(1-\alpha)^{1/n}}{2}$

c)

$n = 20$ $X_{(n)} = 0.48$ $R_\alpha = \{X : X_{(n)} > c\}$

$p(X) = \inf_{\alpha \in (0,1)} \{\alpha : X \in R\} = \inf_{\alpha \in (0,1)} \{\alpha : X_{(n)} > c\}$

$= \inf_{\alpha \in (0,1)} \left\{\alpha : X_{(n)} > \frac{(1-\alpha)^{1/n}}{2}\right\} \Rightarrow \frac{(1-\alpha)^{1/n}}{2} = X_{(n)} \Rightarrow \alpha = 1 - (2X_{(n)})^n$

$\alpha = 1 - (2X_{(n)})^n = 1 - (2 \cdot 0.48)^{20} = 0.558$

2.

$X_1, \dots, X_n \sim N(\mu, \sigma^2)$ $\sigma^2 = 1$ $H_0: \mu = 0$ $H_1: \mu = 1$ $R = \{X : \bar{X}_n > c\}$

a)

$\beta(\mu) = P(\bar{X}_n > c \mid \mu)$ $\bar{X}_n \sim N(\mu, \sigma^2/n)$ $\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} \sim N(0,1)$

$\beta(\mu) = P\left(\frac{\sqrt{n}(\bar{X}_n - \mu)}{\sigma} > \frac{\sqrt{n}(c-\mu)}{\sigma}\right) = 1 - \Phi\left(\frac{\sqrt{n}(c-\mu)}{\sigma}\right) = 1 - \Phi(\sqrt{n}(c-\mu))$

Monotonicity of β:

$\alpha = \beta(0) = 1 - \Phi(\sqrt{n}\, c)$

$c = \frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}}$

b)

$\beta(1) = 1 - \Phi(\sqrt{n}(c-1)) = 1 - \Phi\left[\sqrt{n}\left(\frac{\Phi^{-1}(1-\alpha)}{\sqrt{n}} - 1\right)\right]$

$$= 1 - \phi\left(\phi^{-1}(1-\alpha) - \sqrt{n}\right)$$

c)

$$\lim_{n\to\infty} \beta(0) = \lim_{n\to\infty} \alpha = \alpha$$

$$\lim_{n\to\infty} \beta(1) = \lim_{n\to\infty} 1 - \phi\left(\phi^{-1}(1-\alpha) - \sqrt{n}\right) = 1 - \lim_{n\to\infty} \phi\left(\phi^{-1}(1-\alpha) - \sqrt{n}\right)$$

$$= 1 - \phi(-\infty) = 1 - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-\infty} e^{-t^2/2} dt = 1 - 0 = 1$$

3.

$X_1, \dots, X_n \sim \text{Poisson}(\lambda)$ $\quad H_0 : \lambda = \lambda_0 \quad H_1 : \lambda \neq \lambda_0 \quad \lambda_0 > 0$

a)

$$W = \left| \frac{\hat{\lambda} - \lambda_0}{\hat{se}} \right| > z_{1-\frac{\alpha}{2}}$$

$$L_n(\lambda | X_1, \dots, X_n) = \prod_{i=1}^{n} f(X_i ; \lambda) = \prod_{i=1}^{n} \frac{\lambda^{X_i} e^{-\lambda}}{X_i!}$$

$$\hat{\lambda}_{MLE} = \underset{\lambda}{\operatorname{argmax}}\, L_n(\lambda | X_1, \dots, X_n) = \underset{\lambda}{\operatorname{argmax}}\, \log(L_n)$$

$$\frac{d}{d\lambda} \log(L_n) = 0$$

$$\frac{d}{d\lambda} \log\left(\prod_{i=1}^{n} \frac{\lambda^{X_i} e^{-\lambda}}{X_i!}\right) = \frac{d}{d\lambda} \sum_{i=1}^{n} \log\left(\frac{\lambda^{X_i} e^{-\lambda}}{X_i!}\right) = \frac{d}{d\lambda} \sum_{i=1}^{n} \left(X_i \log(\lambda) - \lambda - \log(X_i!)\right)$$

$$= \frac{d}{d\lambda}\left(\log\lambda \sum_{i=1}^{n} X_i - \lambda n - \sum_{i=1}^{n} \log(X_i!)\right)$$

$$= \frac{1}{\lambda} \sum_{i=1}^{n} X_i - n = 0$$

$$\hat{\lambda}_{MLE} = \frac{1}{n} \sum_{i=1}^{n} X_i \qquad \frac{d^2}{d\lambda^2} \log(L_n) = -\frac{1}{\lambda^2} \sum_{i=1}^{n} X_i < 0$$

$$I(\hat{\lambda}) = \mathbb{E}\left[\left(\frac{\partial \log f(X;\lambda)}{\partial \lambda}\Big|_{\lambda=\hat{\lambda}}\right)^2\right] = \mathbb{E}\left[\left(\frac{\partial}{\partial\lambda} \log\left(\frac{\lambda^X e^{-\lambda}}{X!}\right)\right)^2\right]$$

$$= \mathbb{E}\left[\left(\frac{\partial}{\partial\lambda}(X \log\lambda - \lambda - \log X!)\right)^2\right] = \mathbb{E}\left[\left(\frac{1}{\lambda}X - 1\right)^2\right]$$

$$= \mathbb{E}\left[\left(\frac{X}{\lambda} - \mathbb{E}\left[\frac{X}{\lambda}\right]\right)^2\right] = \operatorname{Var}\left(\frac{X}{\lambda}\right) = \frac{\operatorname{Var}(X)}{\lambda^2} = \frac{\lambda}{\lambda^2} = \frac{1}{\lambda}$$

$$\hat{se} = \frac{1}{\sqrt{n I(\hat{\lambda})}} = \frac{1}{\sqrt{n/\hat{\lambda}}} = \sqrt{\hat{\lambda}/n}$$

$$W = \left| \frac{\frac{1}{n}\sum_{i=1}^{n} X_i - \lambda_0}{\sqrt{\frac{1}{n^2}\sum_{i=1}^{n} X_i}} \right| > z_{1-\frac{\alpha}{2}}$$

```
l0 = 1;
n = 20;
a = 0.05;
m = 10^4;

reject = 0;
for i=1:m
    data = poissrnd(l0, [n 1]);
    W = abs((mean(data)-l0)/(mean(data)/n)^(1/2));
    if W > norminv(1-a/2)
        reject = reject + 1;
    end
end

error = reject/m;
disp(['Estimated type I error rate: ', num2str(error)])

% The estimated type I error rate is right around 0.053
```

```
Estimated type I error rate: 0.0502
```

Published with MATLAB® R2021a

```
A = [7.58 8.52 8.01 7.99 7.93 7.89 7.85 7.82 7.80];
B = [7.85 7.73 8.53 7.40 7.35 7.30 7.27 7.27 7.23];

% Permutation test

K = 10^5;
s = abs(mean(A)-mean(B));

Z = cat(2,A,B);
p = 0;
for i=1:K
    Z_pi = Z(randperm(length(Z)));
    A_pi = Z_pi(1:length(A));
    B_pi = Z_pi(length(A)+1:end);
    s_pi = abs(mean(A_pi)-mean(B_pi));

    if s_pi > s
        p = p + 1;
    end
end

error = p/K;
disp(['Estimated p-value ', num2str(error)])

% The estimated p-value is right around 0.034, which is less than 0.05
 and
% is small engough or us to reject our null hypothesis and say that
 the
% mean soil pH values differ for the two locations.
```

```
Estimated p-value 0.03421
```

Published with MATLAB® R2021a

```
T = [0.225, 0.262, 0.217, 0.240, 0.230, 0.229, 0.235, 0.217];
S = [0.209, 0.205, 0.196, 0.210, 0.202, 0.207, 0.224, 0.223, 0.220,
 0.201];

dmu = mean(T)-mean(S);
se = sqrt(var(T)/length(T)+var(S)/length(S));
W = abs(dmu/se);
p_value = 2 * (1 - normcdf(W));

disp(['p-value ', num2str(p_value)])

% The p-value is 0.0002126, which is less than 0.05 and
% is small engough or us to reject our null hypothesis and say that
 the
% Twain likely did not write the Snodgrass essays.

K = 10^5;
s = abs(mean(T)-mean(S));

Z = cat(2,T,S);
p = 0;
for i=1:K
    Z_pi = Z(randperm(length(Z)));
    T_pi = Z_pi(1:length(T));
    S_pi = Z_pi(length(T)+1:end);
    s_pi = abs(mean(T_pi)-mean(S_pi));

    if s_pi > s
        p = p + 1;
    end
end

error = p/K;
disp(['Estimated p-value ', num2str(error)])

% The estimated p-value is right around 0.00065, which is less than
 0.05 and
% is small engough or us to reject our null hypothesis and say that
 the
% Twain likely did not write the Snodgrass essays.
```

```
p-value 0.0002126
Estimated p-value 0.00084
```

Published with MATLAB® R2021a