Stats PS4 Solution

1. (DGS 9.2.8)

   (a) The PDF's $f_0(x)$ and $f_1(x)$ are as sketched in Fig. ??. Under H_0 it is impossible to obtain a value of X greater than 1, but such values are possible under H_1. Therefore, if a test procedure rejects H_0 only if $x > 1$, then it is impossible to make an error of type I, and $\alpha(\delta) = 0$. Also, $\beta(\delta) = Pr(X < 1|H_1) = 1/2$.

   (b) To have $\alpha(\delta) = 0$, we can include in the critical region only a set of points having probability 0 under H_0. Therefore, only points $x > 1$ can be considered. To minimize $\beta(\delta) = 0$ we should choose this set to have maximum probability under H_1. Therefore, all points $x > 1$ should be used in the critical region.

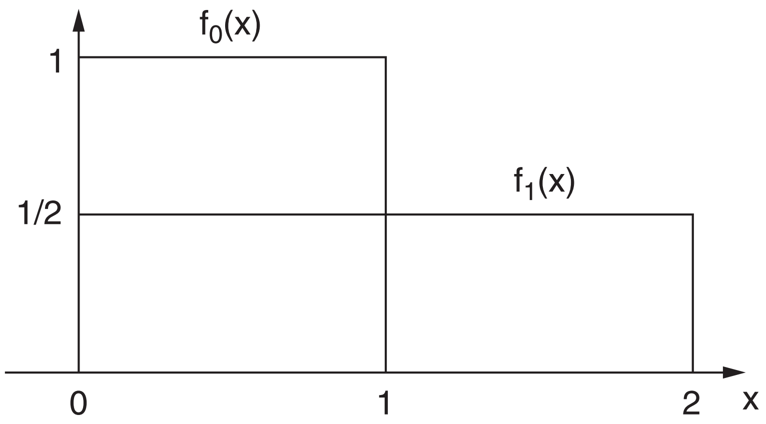

Figure 1.1: Figure for ??.

2. (DGS 9.5.10)
   $\frac{S^2}{\sigma^2} \sim \chi^2_{n-1}$
   Thus, $P\left(S_n^2 \leq c_1\right) = P\left(\chi^2 \leq \frac{c_1}{4}\right) = P\left(S_n^2 \geq c_2\right) = P\left(\chi^2 \geq \frac{c_2}{4}\right) = 0.025$
   Therefore, $\frac{c_1}{4} = \chi^2_{0.975,9} = 2.7$ and $\frac{c_2}{4} = \chi^2_{0.025,9} = 19.023$
   $\Rightarrow c_1 = 10.8, c_2 = 76.095$

3. (DGS 9.6.4)
   Let $Z = \bar{X} - \bar{Y} = \frac{\sigma_1^2}{m} + \frac{\sigma_2^2}{n} = \left(\frac{1}{m} + \frac{k}{n}\right)\sigma_1^2$, thus,

$$\frac{\bar{X_m}-\bar{Y_n}}{\sqrt{\left(\frac{1}{m}+\frac{k}{n}\right)\sigma_1^2}} = Z \sim N(0,1)$$

Also $\sum \frac{(X-\mu_1)^2}{\sigma_1^2} + \sum \frac{(Y-\mu_2)^2}{\sigma_2^2} = \left(S_X^2 + \frac{S_Y^2}{k}\right)\frac{1}{\sigma_1^2} \sim \chi^2_{m+n-2}$.

As we know $\frac{Z}{\sqrt{\chi^2/V}} \sim t$, we have:

$$\frac{(\bar{X_m}-\bar{Y_n})/\sqrt{\left[\left(\frac{1}{m}+\frac{k}{n}\right)\sigma_1^2\right]}}{\sqrt{\left(S_X^2+\frac{S_Y^2}{k}\right)\frac{1}{\sigma_1^2}/(m+n-2)}} = \frac{(m+n-2)^{1/2}(\bar{X_m}-\bar{Y_n})}{\left(\frac{1}{m}+\frac{k}{n}\right)^{1/2}\left(S_X^2+\frac{S_Y^2}{k}\right)^{1/2}} \sim t_{m+n-2}$$

4. (DGS 9.6.6)

   Similarly to the solution to the above Question 9.6.4, we can construct a test statistics:
   $$\frac{(m+n-2)^{1/2}(\bar{X_m}-\bar{Y_n}-\lambda)}{\left(\frac{1}{m}+\frac{1}{n}\right)^{1/2}(S_X^2+S_Y^2)^{1/2}} \sim t_{m+n-2}$$

5. DGS ch. 7.2 problem 10

10. In this exercise

$$f(x|\theta) = \begin{cases} 1 & \text{for } \theta - \frac{1}{2} < x < \theta + \frac{1}{2}, \\ 0 & \text{otherwise,} \end{cases}$$

and

$$\xi(\theta) = \begin{cases} \frac{1}{10} & \text{for } 10 < \theta < 20, \\ 0 & \text{otherwise.} \end{cases}$$

The condition that $\theta - 1/2 < x < \theta + 1/2$ is the same as the condition that $x - 1/2 < \theta < x + 1/2$. Therefore, $f(x \mid \theta)\xi(\theta)$ will be positive only for values of θ which satisfy both the requirement that $x - 1/2 < \theta < x + 1/2$ and the requirement that $10 < \theta < 20$. Since $X = 12$ in this exercise, $f(x \mid \theta)\xi(\theta)$ is positive only for $11.5 < \theta < 12.5$. Furthermore, since $f(x \mid \theta)\xi(\theta)$ is constant over this interval, the posterior p.d.f. $\xi(\theta \mid x)$ will also be constant over this interval. In other words, the posterior distribution of θ must be a uniform distribution on this interval.

6. DGS ch. 7.4 problem 18

18. Suppose that the prior distribution of θ is the Pareto distribution with parameters x_0 and α ($x_0 > 0$ and $\alpha > 0$). Then the prior p.d.f. $\xi(\theta)$ has the form

$$\xi(\theta) \propto 1/\theta^{\alpha+1} \qquad \text{for} \quad \theta \geq x_0.$$

If $X_1, \dots, X_n$ form a random sample from a uniform distribution on the interval $[0, \theta]$, then

$$f_n(\boldsymbol{x} \mid \theta) \propto 1/\theta^n \qquad \text{for} \quad \theta > \max\{x_1, \dots, x_n\}.$$

Hence, the posterior p.d.f. of θ has the form

$$\xi(\theta \mid \boldsymbol{x}) \propto \xi(\theta) f_n(\boldsymbol{x} \mid \theta) \propto 1/\theta^{\alpha+n+1},$$

for $\theta > \max\{x_0, x_1, \dots, x_n\}$, and $\xi(\theta \mid \boldsymbol{x}) = 0$ for $\theta \leq \max\{x_0, x_1, \dots, x_n\}$. This posterior p.d.f. can now be recognized as also being the Pareto distribution with parameters $\alpha + n$ and $\max\{x_0, x_1, \dots, x_n\}$.

7. DGS ch. 7.4 problem 6

6. Suppose that the parameters of the prior gamma distribution of θ are α and β. Then $\mu_0 = \alpha/\beta$. The posterior distribution of θ was given in Theorem 7.3.2. The mean of this posterior distribution is

$$\frac{\alpha + \sum_{i=1}^n X_i}{\beta + n} = \frac{\beta}{\beta + n}\mu_0 + \frac{n}{\beta + n}\overline{X}_n.$$

Hence, $\gamma_n = n/(\beta + n)$ and $\gamma_n \to 1$ *as* $n \to \infty$.

$\gamma_n \sim \frac{n}{n}$

8. DGS ch. 7.4 problem 12

12. (a) A's prior distribution for θ is the beta distribution with parameters $\alpha = 2$ and $\beta = 1$. Therefore, A's posterior distribution for θ is the beta distribution with parameters $2+710 = 712$ and $1+290 = 291$. B's prior distribution for θ is a beta distribution with parameters $\alpha = 4$ and $\beta = 1$. Therefore, B's posterior distribution for θ is the beta distribution with parameters $4 + 710 = 714$ and $1 + 290 = 291$.

(b) A's Bayes estimate of θ is $712/(712+291) = 712/1003$. B's Bayes estimate of θ is $714/(714+291) = 714/1005$.

(c) If y denotes the number in the sample who were in favor of the proposition, then A's posterior distribution for θ will be the beta distribution with parameters $2 + y$ and $1 + 1000 - y = 1001 - y$, and B's posterior distribution will be a beta distribution with parameters $4+y$ and $1+1000-y = 1001 - y$. Therefore, A's Bayes estimate of θ will be $(2 + y)/1003$ and B's Bayes estimate of θ will be $(4 + y)/1005$. But

$$\left|\frac{4+y}{1005} - \frac{2+y}{1003}\right| = \frac{2(1001-y)}{(1005)(1003)}.$$

This difference is a maximum when $y = 0$, but even then its value is only

$$\frac{2(1001)}{(1005)(1003)} < \frac{2}{1000}.$$