MAST20005/MAST90058: Week 7 Problems

Some useful information for many of the problems is shown at end of this problem sheet.

1. A ball is drawn from one of two bowls. Bowl A contains 100 red balls and 200 white balls; Bowl B contains 200 red balls and 100 white balls. Let p denote the probability of drawing a red ball from the chosen bowl. Then p is unknown as we don't know which bowl is being used. We shall test the simple null hypothesis, $H_0\colon p = 1/3$, against the simple alternative, $H_1\colon p = 2/3$. We draw three balls at random with replacement from the selected bowl. Let X be the number of red balls drawn. Let the critical region be $x \in \{2, 3\}$. What are the probabilities α and β respectively of Type I and Type II errors?

2. Let $Y \sim \mathrm{Bi}(100, p)$. To test $H_0\colon p = 0.08$ against $H_1\colon p < 0.08$, we reject H_0 if $Y \leqslant 6$.

   (a) Determine the significance level α of the test.

   (b) Find the probability of a Type II error if, in fact, $p = 0.04$.

3. If a newborn baby has a birth weight that is less than 2500 grams we say the baby has a low birth weight. The proportion of babies with low birth weight is an indicator of nutrition for the mothers. In the USA, approximately 7% of babies have a low birth weight. Let p be the proportion of babies born in the Sudan with low birth weight. Test the null hypothesis $H_0\colon p = 0.07$ against the alternative $H_1\colon p > 0.07$. If $y = 23$ babies out of a random sample of $n = 209$ babies had low birth weight, what is your conclusion at the following significance levels:

   (a) $\alpha = 0.05$?

   (b) $\alpha = 0.01$?

   (c) Find the p-value of this test. (Recall the p-value is the probability of the observed value or something more extreme when the null hypothesis is true).

4. Let p_m and p_f be the respective proportions of male and female white crowned sparrows that return to their hatching site. Give the endpoints for a 95% confidence interval for $p_m - p_f$, given that 124 out of 894 males and 70 out of 700 females returned. Does this agree with the conclusion of the test of $H_0\colon p_m = p_f$ against $H_1\colon p_m \neq p_f$ with $\alpha = 0.05$?

5. Vitamin B_6 is one of the vitamins in a multivitamin pill manufactured by a pharmaceutical company. The pills are produced with a mean of 50 milligrams of vitamin B_6 per pill. The company believes there is a deterioration of 1 milligram per month, so that after 3 months they expect that $\mu = 47$. A consumer group suspects that $\mu < 47$ after 3 months.

   (a) Define a critical region to test $H_0\colon \mu = 47$ against $H_1\colon \mu < 47$ with a significance level of $\alpha = 0.05$ based on a random sample of size $n = 20$ and assuming a normal distribution.

   (b) If the 20 pills resulted in a mean of $\bar{x} = 46.94$ and a standard deviation of $s = 0.15$, what is your conclusion?

   (c) Give bounds for the p-value of this test.

6. Let X represent the height of professional female volleyball players. Assume that $X \sim \mathrm{N}(\mu, \sigma^2)$ approximately. Suppose it is known that $\mu = 1.9$ metres worldwide. Do Australian female players differ from this? We explore this using a random sample of size $n = 9$.

(a) Define the null hypothesis.
(b) Define the alternative hypothesis.
(c) Define a critical region for which $\alpha = 0.05$.
(d) Calculate the value of the test statistic if $\bar{x} = 2.05$ and $s = 0.17$.
(e) What is your conclusion?
(f) Give bounds for the p-value of this test.

7. In May, the weights of 2-kilogram boxes of laundry detergent had a mean of 2.13 kilograms with a standard deviation of 0.095. The goal was to decrease the standard deviation. The company decided to adjust the filling machines and then test $H_0\colon \sigma = 0.095$ against $H_1\colon \sigma < 0.095$. In June, a random sample of size $n = 20$ gave $\bar{x} = 2.10$ and $s = 0.065$.

   (a) At an $\alpha = 0.05$ significance level, was the company successful?
   (b) What is an approximate p-value for this test?

8. The World Health Organisation collects air quality data around the world, which includes a measurement of suspended particles in μ g/m^3. Let X and Y equal the concentration of suspended particles in the city centres of Melbourne and Sydney, respectively. Using $n = 13$ observations of X and $m = 16$ observations of Y, we shall test $H_0\colon \mu_X = \mu_Y$ against $H_1\colon \mu_X \neq \mu_Y$ using a significance level of 5%.

   (a) Define the test statistic and the critical region assuming the variances are equal.
   (b) If $\bar{x} = 72.9$, $s_X = 25.6$, $\bar{y} = 81.7$ and $s_Y = 28.3$, calculate the value of the test statistic and state your conclusion.

   (c) Give bounds for the p-value of this test.
   (d) Test whether the assumption of equal variances is valid. Let $\alpha = 0.05$.

Some potentially helpful R output:


```
> dbinom(0:3, 3, 1/3)
[1] 0.29629630 0.44444444 0.22222222 0.03703704
> dbinom(0:3, 3, 2/3)
[1] 0.03703704 0.22222222 0.44444444 0.29629630
> pnorm(c(-0.737, -0.553, 1.276, 1.531, 2.269))
[1] 0.2305612 0.2901317 0.8990222 0.9371153 0.9883658

> p1 <- c(0.75, 0.9, 0.95, 0.975, 0.99)
> qnorm(p1)
[1] 0.6744898 1.2815516 1.6448536 1.9599640 2.3263479
> qt(p1, 27)
[1] 0.683685 1.313703 1.703288 2.051831 2.472660
> qt(p1, 20)
[1] 0.6869545 1.3253407 1.7247182 2.0859634 2.5279770
> qt(p1, 19)
[1] 0.6876215 1.3277282 1.7291328 2.0930241 2.5394832
> qt(p1, 8)
[1] 0.7063866 1.3968153 1.8595480 2.3060041 2.8964594

> p2 <- c(0.025, 0.05, 0.95, 0.975)
> qchisq(p2, 19)
[1]  8.906516 10.117013 30.143527 32.852327
> qf(p2, 12, 15)
[1] 0.3147424 0.3821387 2.4753130 2.9632824
```