Stat 21 Homework 3

Problem 5 Solutions

Problem 5

After losing several times in a street performance game, you suspect that the die used by the performer may be unfair. To check, you roll the die 60 time, recording the number of times each face appears. Do these results case doubt on the die's fairness? If the die is fair, how many times would you expect each face to occur?

Face	Count
1	11
2	8
3	9
4	15
5	10
6	7

To answer this question, perform a chi-squared goodness-of-fit test. Clearly state the null and alternative hypotheses, check the necessary conditions, identify the degrees of freedom and report the p-value.

Solution Problem 5:

```
obs_counts <- c(11,8,9,15,10,7)
probs <- rep(1/6, 6)
chisq.test(obs_counts, p=probs)
```

```
##
##  Chi-squared test for given probabilities
##
## data:  obs_counts
## X-squared = 4, df = 5, p-value = 0.5494
```

Alternatively, students may have solved the problem "by hand" with something such as

```
obs_counts <- c(11,8,9,15,10,7)
exp_counts <- rep(10,6)
test_stat = sum((obs_counts-exp_counts)^2/exp_counts)
pchisq(test_stat, df=5, lower.tail=FALSE)
```

```
## [1] 0.549416
```

H_0 : The dice is fair.

or

H_0 : The probability of landing on any of the six faces is the same.

H_A : The dice is loaded.

or

H_0 : The probability of landing on each of the six faces are not all the same.

Degrees of freedom = 5

Significance level: $\alpha = 0.05$

P-value = 0.55

Assumptions:

1) The observed sample is representative of the population and independently drawn. Although this is not a simple random sample, this condition is satisfied because each roll of the die is an independent simulation of the random variable.

2) The expected cell counts are all at least 5. This is true since all expected cell counts are 10.

Conclusion: With a p-value this large, we fail to reject the null hypothesis (at any typical significance level for $\alpha \leq 0.10$.) This means that based on our sample of data, we do not have any statistical evidence that this dice is loaded.

Notes on grading:

The most important part of this question is the conclusion. To receive full credit, answers to this question should contain each of components mentioned above and provide R code. The conclusion and the interpretation of the conclusion in the context of the problem should be worth the most points for this problem however.