Probability and Statistics for JMC
Exercises 7 — Hypothesis Testing

1. To decide from which manufacturer to purchase 2000 PCs for its undergraduates, a university decided to carry out a test. It bought 50 PCs from manufacturer A and 50 from manufacturer B. Over the course of a six month period, 6 of those from manufacturer A experienced problems, as did 10 from manufacturer B. This suggested that, to be on the safe side, the university should buy machines from manufacturer A. However, A's machines are more expensive than B's. Before making the decision to buy from A, the university wanted to be confident that the difference was a real one, and was not merely due to chance fluctuation. Carry out a test to investigate this.

   You may wish to use this extract from the chi-squared table:

Degrees of freedom	Upper tail area		
	.10	.05	.01
1	2.706	3.841	6.635
2	4.605	5.991	9.210
3	6.251	7.815	11.345
4	7.779	9.488	13.277
5	9.236	11.070	15.086

2. In an ESP experiment, a subject in one room is asked to state the colors of 100 cards chosen by someone in another room. The chooser has a pack of 25 red and 25 blue cards and selects the cards at random, with replacement. If the subject gets 64 right, determine whether the results are significant at the 1% level. Clearly write down your null hypothesis, alternative hypothesis, the test statistic you use, the rejection region, and interpret your conclusions.

   Hint: you may regard 100 as a large sample. You may wish to use this table giving quantile function of the standard normal distribution for various upper tail areas.

Upper tail area α	0.1	.05	.025	.01	.005	.001
$\Phi^{-1}(1-\alpha)$	1.28	1.64	1.96	2.33	2.58	3.09

3. Charles Darwin measured differences in height for 15 pairs of plants of the species *Zea mays*. (Each plant had parents grown from the same seed – one plant in each pair was the progeny of a cross-fertilisation, the other of a self-fertilisation. Darwin's measurements were the differences in height between cross-fertilised and self-fertilised progeny.) The data, measured in eighths of an inch, are given below.

$$49, \quad -67, \quad 8, \quad 16, \quad 6, \quad 23, \quad 28, \quad 41, \quad 14, \quad 29, \quad 56, \quad 24, \quad 75, \quad 60, \quad -48$$

(a) Supposing that the observed differences $\{d_i \mid i = 1, \ldots, 15\}$ are independent observations of a normally distributed random variable D with mean μ and variance σ^2, state appropriate null and alternative hypotheses for a two-sided test of the hypothesis that there is no difference between the heights of progeny of a cross-fertilised and self-fertilised plant, and state the null distribution of an appropriate test statistic.

(b) Obtain the form of the rejection region for the test you defined in part (a), assuming a 10% significance level.

(c) Calculate the value of the test statistic for this data set, and state the conclusions of your test.

You may want to use the following extract from a t-distribution, giving the point x corresponding to specified areas under the upper tail of a t-distribution with df degrees of freedom.

df	5%	10%
13	1.7709	1.3502
14	1.7613	1.3450
15	1.7531	1.3406

4. Some of "Student's" original experiments involved counting the numbers of yeast cells found on a microscope slide. The results of one experiment are given in the table below, which shows the number of small squares on a slide which contain 0, 1, 2, 3, 4, or 5 cells. We want to use these data to see if the mean number of cells per square is 0.6, using the 5% significance level.

Number of cells in square	0	1	2	3	4	5
Frequency	213	128	37	18	3	1

(a) State the null hypothesis and the alternative hypothesis.

(b) The distribution of the numbers of cells is far from normal; it can take only positive integer values, it is very far from symmetric, and dies away very quickly. Which familiar distribution might be appropriate as a model for these data?

(c) Estimate the mean and the variance of the distribution you suggested in part (b).

(d) Explain why a critical region with the form

$$\left\{\bar{x} < \mu - 1.96\sqrt{\frac{\mu}{n}}\right\} \cup \left\{\bar{x} > \mu + 1.96\sqrt{\frac{\mu}{n}}\right\}$$

would be a reasonable region, making sure you explain the 1.96, the term $\sqrt{\dfrac{\mu}{n}}$ and the implications of the union. What is the test statistic in mind?

(e) Compute the limits of the critical region.

(f) Draw a conclusion about whether or not the null hypothesis can be rejected.

5. A survey of 320 families with 5 children each, gave the distribution shown below. Is this table consistent with the hypothesis that male and female children are equally probable? Obtain results for both the 1% and 5% levels. Work through the details of the test – don't just hit the chi-squared button on a statistical calculator.

[You may wish to use the table extract from Q1.]

boys : girls	5 : 0	4 : 1	3 : 2	2 : 3	1 : 4	0 : 5
Number of families	18	56	110	88	40	8

6. As part of a telephone interview, a sample of 500 executives and a sample of 250 MBA students were asked to respond to the question "Should corporations become more directly involved with social problems such as homelessness, education, and drugs?" The results are shown below. Test the hypothesis that the patterns of response for the two groups are the same.

[You may wish to use the table extract from Q1.]

	More involved	Not more involved	Not sure
Executives	345	135	20
MBA students	222	20	8

7. (a) For a test at a fixed significance level, and with given null and alternative hypotheses, what will happen to the power as the sample size increases?

(b) For a test of a given null hypothesis against a given alternative hypothesis, and with a fixed sample size, describe what would happen to the power of the test if the significance level was changed from 5% to 1%.

(c) A test of a given null hypothesis against a given alternative hypothesis, with a sample of size n and significance level of α, has a power of 80%. What change could I make to the test to increase my chance of rejecting a false null hypothesis?

(d) How can we attain a test which has a very low probability of Type I error and also a very low probability of Type II error?

8. The data below show the frequency with which each of the balls numbered 1 to 49 have appeared in the main draw in the National Lottery between its inception in November 1994 until December 2008.

   This table shows that there are substantial differences between the numbers of times different balls have appeared — for an extreme example, the number 20 has been drawn just 134 times, whereas 38 has appeared 197 times, almost 50% more often.

   Should we conclude from this table that the balls have different probabilities of appearing?

Ball	Freq.	Ball	Freq.	Ball	Freq.	Ball	Freq.	Ball	Freq.
1	160	11	185	21	148	31	180	41	137
2	168	12	178	22	169	32	170	42	167
3	163	13	147	23	183	33	176	43	182
4	163	14	160	24	164	34	159	44	184
5	155	15	157	25	182	35	173	45	162
6	174	16	147	26	161	36	148	46	164
7	168	17	158	27	173	37	150	47	177
8	158	18	164	28	165	38	197	48	177
9	176	19	166	29	164	39	166	49	169
10	171	20	134	30	177	40	178		

   [To save you a lot of data entry you might use the following way of rewritting a particular polynomial:

$$\sum_{i=1}^{49}(n_i - x)^2 = 49x^2 - 16308x + 1364626,$$

   where n_i is the frequency of ball i in the above table.]

Partial answers:

1. $\chi^2 = 1.19$
2. Test statistic $z = 2.8$ (normal approx) or can use binomial directly to get a more exact test.
3. (a) $H_0 : \mu = 0$ vs. H_1 :??; null distribution is Student-t(14). (b) $R = \{t : |t| > 1.761\}$. (c) $t = 2.15$
4. (a) $H_0 : \mu = 0.6$ vs. H_1 :??; (c) mean=var=0.6825; (e) $0.5241, 0.6759$; (f) reject
5. $\chi^2 = 11.96$, reject at 5% but not at 1%
6. $\chi^2 = 38.04$, 2 dof
8. $\chi^2 = 46.475$, 48 dof