1. Charles Darwin conducted an experiment to determine whether seedlings from cross–fertilized plants grow better than those from self–fertilized plants. He chose fifteen plants, and covered them with netting so that insects could not fertilize them. He fertilized some flowers on each plant with their own pollen, and an equal number of flowers on each plant with pollen from another plant. The seeds from the flowers were ripened and placed in sand to germinate. Darwin placed two seedlings of the same age from each plant in a pot, one from a seed from a self–fertilized flower, the other from a cross–fertilized flower. The heights (in inches) of the resulting plants were recorded after two months.
   In analysing the data produced by Darwin, which method could sensibly be used? (choose all that apply):

   (a) One sample t test

   (b) Two–sample t test

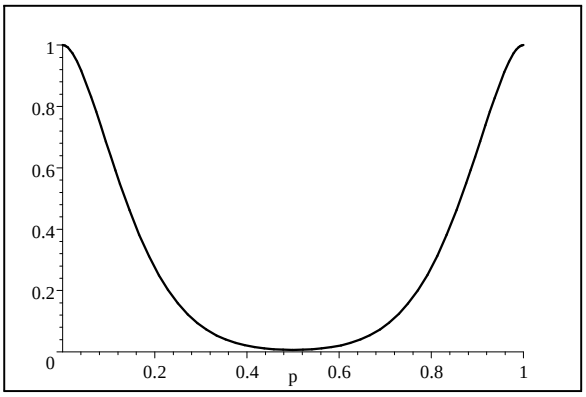
   (c) **Paired t test**

   (d) Kruskal–Wallis test

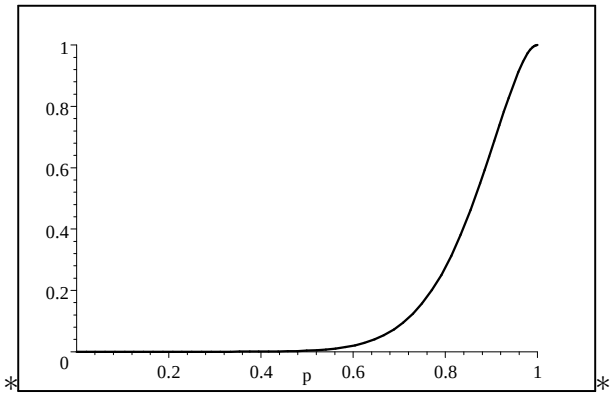
   (e) **Signed Rank test**

2. A friend claims to have the ability to predict the outcome of the toss of a fair coin. You devise a test for this claim, in which an independent person will toss a coin twelve times, and prior to each toss your friend will attempt to predict the outcome. In creating a hypothesis test for the data gathered, your parameter p is the probability your friend can predict the outcome of a fair coin toss. Taking the number of correct guesses as the test statistic, you decide to reject your null hypothesis if your friend makes eleven or twelve correct guesses. The power function for this test is which of the following?

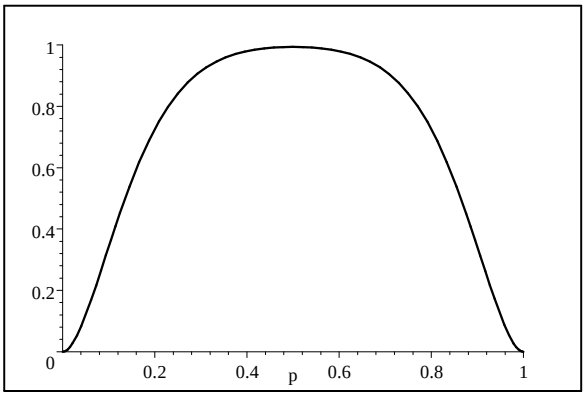
(a)


1
0.8
0.6
0.4
0.2
0
0.2
0.4
p
0.6
0.8
1

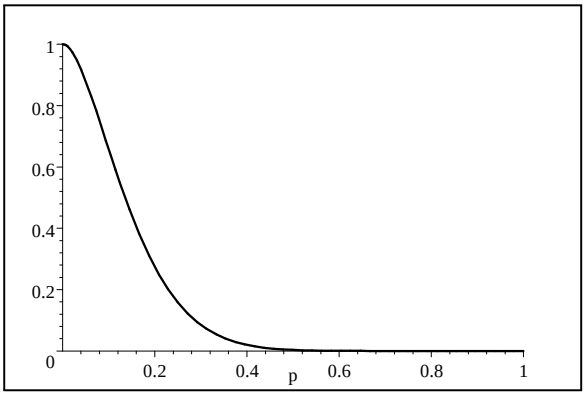
(b)


1
0.8
0.6
0.4
0.2
0
0.2
0.4
p
0.6
0.8
1

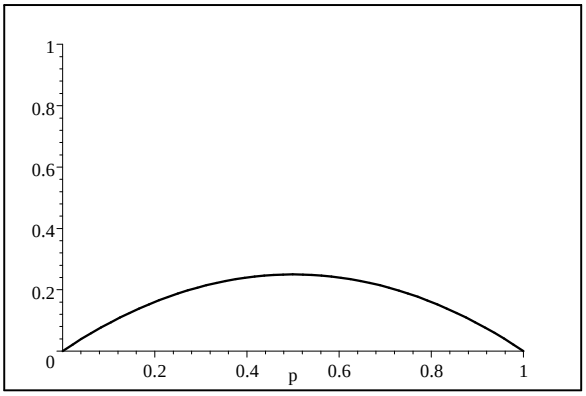
(c)


1
0.8
0.6
0.4
0.2
0
0.2
0.4
p
0.6
0.8
1

(d)



(e)



3. A market research company wonders if customer satisfaction might vary with gender. They will administer a questionnaire to a random sample of 500 customers, on which customers will report their gender (female or male) and their level of satisfaction (low, medium, or high). They will use a chi-squared test with the null hypothesis being independence of gender and level of satisfaction. This will involve comparing the χ^2 test statistic to the χ^2_d distribution, where the degrees of freedom is:

   (a) $d = 1$

   (b) $*d = 2*$

   (c) $d = 3$

   (d) $d = 4$

(e) $d = 5$

Since there are two levels of gender and three levels of satisfaction, cross classification will yield $m = 6$ categories (a 2 by 3 table, or 3 by 2 table, whichever you prefer). Under the null hypothesis that gender and satisfaction are independent, there are $q = 3$ free parameters all together (two to describe the distribution of customers across the three levels of satisfaction plus one to describe the distribution of customers across the two levels of gender). Hence

$$d = (m - 1) - q = 5 - 3 = 2.$$

4. A study aimed to test the response to a new production technique in car manufacturing. Two groups of ten fitters were shown the new technique, while two control groups of ten fitters continued with the standard method. There were insufficient fitters in a single factory to conduct the experiment, so one of each of the two treatment groups was in factory A and the other in factory B. The twenty fitters at each factory were assigned at random to either the new technique or the control group. The responses of all individuals were recorded, these being the times (in minutes) taken to complete the fitting task to a certain specification. The resulting ANOVA statistic proved to be significant, rejecting the null hypothesis that the mean responses were the same, the means in the groups using the standard method being somewhat higher than those in control groups. The study included which of the following (choose all that apply):

   (a) Blinding
   (b) **Randomization**
   (c) Missing values
   (d) A placebo
   (e) **Blocking**

5. Dependent on birth date, each person is assigned to one astrological "sign". There are twelve such signs, commonly known as the "Signs of the Zodiac". A study investigated a relationship between astrological

sign and heart rate (measured in bpm (beats per minute)). The study asked each subject for their astrological sign and also had their heart rates recorded while at rest. To test the null hypothesis that the expected heart rate is the same for people of each astrological sign, an ANOVA table was computed. The value of the test statistic was 0.61, to be compared with the $F_{11,813}$ distribution. Based on this information, you could deduce which of the following?

(a) The mean heart rate was 813 bpm.

(b) The variance of all the heart rates recorded is 0.61 bpm^2.

(c) There were eleven people in the study born under each of the twelve astrological signs.

(d) There is 61% of the variation in the heart rates that is accounted for by astrological sign.

(e) **There were 825 subjects in the study.**

6. A gardener wished to investigate which type of tomato plant, either A or B say, produces the best yield in his greenhouse. He planted three plants of each type, and measured the yields (in kg) from each. The data are given below:

Plant type	
A	B
2.6, 2.2, 1.9	3.6, 3.0, 2.3

(a) In constructing a test for this experiment, what would be your null hypothesis?
The null hypothesis could be that the distributions of tomato yields are the same for both plant types. Alternatively, one might test whether the underlying means (or medians) are equal.

(b) In conducting a hypothesis test here, would you take a one–sided or two–sided alternative? Explain your reasoning.
A two–sided alternative seems reasonable here – there is no suggestion that the gardener believed one type produced greater yield before conducting the study and both one–sided alternatives are viable.

(c) We can construct a test based on choosing a suitable test statistic and considering the number of regroupings of the data that lead to a test statistic at least as extreme as the one observed. By regroupings, we mean the ways of labelling the six observations so that three are from plant type A and three are from type B.
In how many ways can the data be so regrouped?
Six observations can be grouped into two sets of three in

$$\binom{6}{3} = 20$$

ways.

(d) We will take as our test statistic the difference in the group means. In how many ways can the data be regrouped to provide a test statistic at least as inconsistent with the null hypothesis as the data observed?
The data observed allocate $(2.6, 2.2, 1.9)$ *to plant A, and the test statistic is –0.734 kg. Since we use a two–sided test, swapping the group of each datum gives test statistic 0.734 kg, this being as inconsistent with the null hypothesis. More inconsistent would be allocating* $(2.3, 2.2, 1.9)$ *to plant A, as is having* $(3.6, 3.2, 2.6)$ *being from A (these cases have test statistic* ± 0.934*). So four regroupings are at least as inconsistent with the null hypothesis as the one observed (three if we exclude the case observed).*

(e) Compute the p–value for the test. What would you conclude about the null hypothesis?
We see there are 4 of the 20 possible groupings giving a test statistic at least as inconsistent as that observed. This yields a p–value of

$$\frac{4}{20} = 0.2.$$

Hence there is insufficient evidence to reject the null hypothesis that the type plant types produce yield with the same mean.

7. A study will be carried out to assess the hypothesis that all the numbers from 1 to 49 are equally likely to be drawn as winning numbers in Lotto 6/49. The study will track the 3744 winning numbers drawn over the course of three years (there are bi-weekly draws of 6 numbers at a time).

A chi-squared statistic will be constructed to compare the 49 observed frequencies to appropriate expected counts.

(a) What are the expected counts under the null hypothesis?
Under the null hypothesis, each number has a 1/49 chance of being selected on each draw. Hence each of the expected counts is just $3744/49 = 76.4$.

(b) The chi-squared test statistic will be compared to the χ^2_d distribution. What value of the degrees-of-freedom d should be used? Explain your reasoning.
We are basing our chi-squared test on the observed counts for $m = 49$ *categories. Note that in this problem we do not have to estimate anything from the data to get the expected counts, i.e., there are* $q = 0$ *free parameters under the null hypothesis. Thus we would be comparing our test statistic to the chi-squared distribution with* $d = (m - 1) - q = 48$ *degrees of freedom.*

8. The following data are the heights, in metres, of twelve plants grown in an agricultural experiment:

$$\begin{array}{cccccc} 7.60, & 5.80, & 6.82, & 4.86, & 5.65, & 6.22, \\ 4.79, & 4.35, & 5.56, & 5.24, & 3.82, & 5.50. \end{array}$$

(a) We wish investigate whether the above data appear to be consistent with arising from a Normal distribution. Describe what probability plot you would create, defining any terminology you require.
We sort the data into order, smallest to largest, and plot against the corresponding Normal scores, these being the expected values of an ordered sample of size twelve from $N(0, 1)$.

(b) How would you expect your plot to appear if the data arise from a Normal distribution with mean 5 kg and standard deviation 0.5 kg?
We would expect the plot to be approximately linear with slope 0.5 and intercept 5.

(c) Sketch how would you might expect your plot to appear if the data arose from a distribution that is right skewed.
The plot could resemble something like the following: