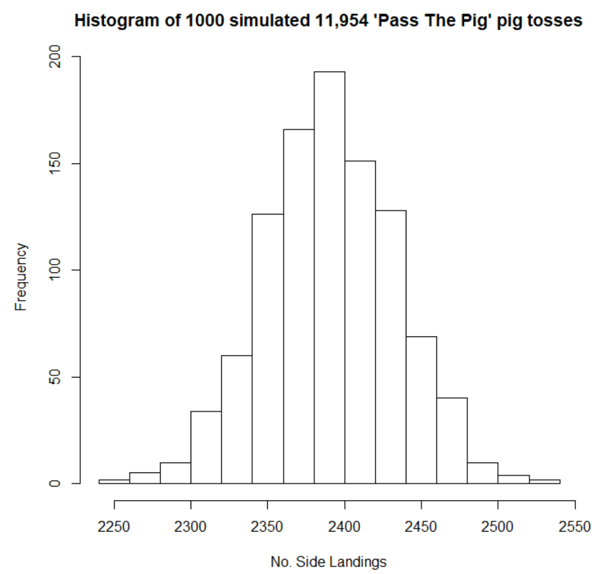
1. In the game "Pass the Pigs" players toss plastic pigs and score points depending on how they land. An image of the pigs is shown below.

By Larry D. Moore, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=28974299

For a single pig, there are five possible ways the pigs can land when tossed: the pig can land on its back, side, upright on its legs, snout, or resting on its snout and ear. You are interested in exploring the relative probabilities for the outcomes. Suppose the parameter of interest is the probability p a pig lands on its side on a single toss. We will take the *null hypothesis* is that each of the five ways of landing are equally likely.

   (a) When tossing a single pig in the game, a player scores no points if the pig lands on its side but will score points for the other outcomes. Given that games of chance tend to reward players for outcomes that are rare, state your *alternative hypothesis* about the parameter of interest. That is, state what you would take as the alternative to the null hypothesis. Explain your choice clearly. *Given that no points are assigned for what might be assumed the* *most likely outcome, a one–sided alternative seems reasonable, being* $p > 1/5$.

(b) Kern (2006) reports a study that involved tossing a single pig from the game 11,954 times. Of those tosses, the pig landed on its side 7782 times. We explore the data by computer–based simulation assuming the null hypothesis. The results from 1000 simulations of 11,954 pig tosses are shown below, each one simulated under the assumption from the null hypothesis. The number of times the pig landed on its side was recorded for each simulation.



Looking at the above, which of the following is the percentage of simulated samples that are at least as inconsistent with the null hypothesis as the data from the study?

i. **0%**
ii. 33%
iii. 50%
iv. 66.7%
v. 99%

(c) State clearly in a sentence what you conclude from the study.
If the pigs are equally likely to fall on their sides as any other outcome, the data observed in the study are extremely unlikely.

(d) Using the data from the Kern study, create a 95% confidence interval for the probability a pig in the game lands on its side on a single toss during the game "Pass The Pigs". Show your working clearly.
The point estimate is $\hat{p} = 7782/11954$, *the estimated standard deviation of the estimate being*

$$\sqrt{\frac{7782/11954\,(1 - 7782/11954)}{11954}} = 0.004\,36.$$

Hence a 95% confidence interval is $7782/11954 \pm 2\times\ 0.004\,36 = (0.642\,, 0.660)$.

(e) Which value of p would give the wider confidence interval: the actual value or the value under the null hypothesis? Explain your reasoning clearly.
The true proportion appears to be around 0.65 and this is closer to 0.5 than the value under the null hypothesis (0.2). We recall that the s.d. of the sample proportion is largest when the population proportion is 0.5. Hence the actual proportion here would give rise to wider confidence intervals than if it were the value under the null hypothesis.
Kern, J. C. (2006): Pig Data and Bayesian Inference on Multinomial Probabilities. *Journal of Statistics Education* **14** (3).
`doi:10.1080/10691898.2006.11910593`

2. Twelve different cars were tested to find their mileage with 10 gallons of regular fuel (R) and their mileage with the same quantity of fuel with an additive (A). The data collected

	Car											
	1	2	3	4	5	6	7	8	9	10	11	12
Mileage on R:	152	166	137	147	136	140	130	163	160	122	147	124
Mileage on A:	154	158	150	154	130	145	147	160	168	132	151	136

(a) In this context, what is the parameter of interest?
The mean difference in mileage for cars between 10 gallons of the regular fuel and 10 gallons of the fuel with an additive.

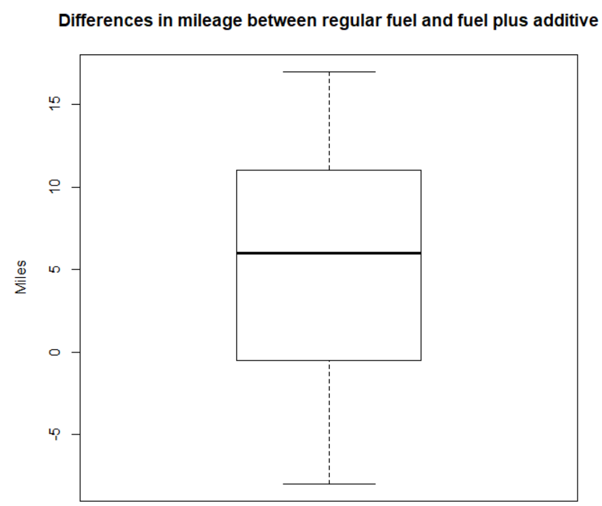
(b) If considering a hypothesis test in this context, which of the following is a statement of the null hypothesis being tested?

   i. The proportion of cars that will run for longer on 10 gallons of fuel plus the additive compared to 10 regular gallons is 0.5.
   ii. The difference in the sample means between the mileages on the regular fuel and the fuel with the additive is zero.
   iii. **The mean difference in mileage is zero over all cars between using 10 gallons of the regular fuel and 10 gallons of the fuel plus the additive.**
   iv. The mean mileage for all cars on 10 gallons of regular fuel plus the additive is higher than the mean mileage using the regular fuel.
   v. Any car will have the same mileage running on 10 gallons of the regular fuel or 10 gallons of the regular fuel plus the additive.

(c) This is an example of a matched pairs design. The researchers instead could have taken a sample of cars and run them on 10 gallons of regular fuel and compared with another sample of cars running on the regular fuel plus the additive. Why is the matched pairs design, where both mileages were taken on each car, likely to be better at addressing the research question? It is because the matched pairs design ...

   i. produces more data.
   ii. only involves differences within individual cars, which is likely to have less variation than differences between different cars.
   iii. enables us to compute the correlation between the mileages both with and without the fuel additive.
   iv. is less expensive to run as it requires fewer cars.
   v. reduces the risk of measurement error.

(d) Below is a boxplot of the differences in the mileages for the twelve cars across the two types of fuel.



Do you think that the Central Limit Theorem is good to apply here? Explain your thinking in a sentence (or two).
The sample size is quite small but the distribution of the differences appears quite symmetric. It is hard to "break" the CLT even when the sample size is quite small unless the population distribution is very skewed.

(e) The mean of the differences for the data above is 5.08 miles and the standard deviation is 7.74 miles. Assuming the null hypothesis holds, sketch the approximate distribution governing how the sample mean would vary across many repeats of the experiment. Indicate the approximate location of the sample mean on your picture.
By CLT, the distribution of the sample mean will be approximately Normal, centred at zero and with standard deviation about

$$\frac{7.74}{\sqrt{12}} = 2.234$$

miles. The sample mean is therefore more than two standard deviations above zero, so is quite high in the upper tail.

(f) Use the data to find a 95% confidence interval for the parameter of interest.

The interval is $5.08 \pm 2 \times 2.234 = (0.612, 9.548)$ *miles.*

(g) In the context of repeatedly sampling twelve cars and running them on the two types of fuel, interpret your 95% confidence interval clearly.

If repeatedly taking samples of 12 cars, finding the difference in mileage on each running on both fuel types, and computing the interval for each sample as above, about 95% of such intervals would contain the true mean difference in mileage between the two fuel types.

(h) Write a sentence (or two) to communicate your conclusion about the impact of the additive on car mileage.

The data appear quite unlikely if in fact the additive has no impact on mileage. In fact, it appears likely the additive has a positive effect on miles per gallon.

3. *Racial steering* occurs when a real estate agent tends to show prospective property renters only homes in neighbourhoods already dominated by the race of the prospective renter. Such action is against the law in some countries. Connor and Kadane (2001) write about an American court case where it was alleged that a real estate agent was applying racial steering to renters. Data were provided about how the race of a potential renter of an apartment in a complex appeared to influence the section in the complex the realtor showed the potential renter. The complex split into two sections: section A that was predominantly white and B that was predominantly black. Suppose the data over a two–year period were as below:

		Potential Renter Race	
		White	Black
Section shown	A	81	14
	B	83	34

Connor, J.T. and Kadane, J.B. (2001): Alleged racial steering in an apartment complex. *Chance* **14**, No.2, 19-22.

(a) What is the parameter of interest in this study?

i. The proportion of white potential renters shown an apartment in section A.
ii. The difference in the proportions of white and black potential renters in the sample shown section A.
iii. **The difference in the probabilities of white and black potential renters being shown section A by the realtor.**
iv. The difference in the counts of white and black potential renters in the sample who were shown section A.
v. Whether the realtor was racist.

(b) Based on the data provided, what is your estimate of the parameter of interest (to three decimal places)?
The estimate would be the difference in the sample proportions,

$$\frac{81}{164} - \frac{14}{48} = 0.202$$

(*or alternatively* -0.202)

(c) In testing a hypothesis about the parameter of interest, what would your null hypothesis be?

i. There was no difference between the number of white and black potential renters.
ii. There was no difference in the proportions of white and black potential renters wanting an apartment in section A.
iii. White potential renters were just as likely to rent an apartment in the complex as black potential renters.
iv. **The chance of being shown an apartment in section A did not depend on the race of the potential renter.**
v. There was no difference in the numbers of apartments being available to rent in the two sections of the complex.

(d) You would take the alternative hypothesis to be

i. two–sided.
ii. one–sided, left–tailed.
iii. **one–sided, right–tailed.**

iv. either a one–sided or two–sided alternative, as it does not matter.

Justify your choice clearly:

The data were presented in a court case for racial steering, a kind of discrimination in which the race of a potential renter influences the location of apartments they are shown by a realtor. It is inconceivable that such a case would go to court if in fact the realtor tended to show potential renters apartments in the section dominated by residents of a ***different*** *race. So a one–sided alternative hypothesis seems sensible.*

(Left–tailed alternative is appropriate if difference (in proportions) defined with opposite sign.)

(e) Suppose you were to conduct an exploration for these data using a *physical* simulation. Explain clearly (i) what materials you would use and (ii) exactly how you would conduct your simulation–based exploration. You can assume you have a lot of time to do your study. You should indicate how you would form your conclusion from your exploration.

i. Materials required:

We would require cards of different colours: 164 "white" and 48 "black". (Actual colours of cards is irrelevant)

ii. What you would do:

Shuffle the 212 cards together. Then deal out the cards into two sets, one for section A with 95 cards, the other for section B with 117 cards. Find the proportion of white cards in section A and the proportion of black cards in section A. Find the difference in the two proportions. Repeat these steps lots of times, say 100. Find the proportion of times the difference in proportions is at least as big as 0.202. Then decide how surprising the sample difference of 0.202 appears if there were no association between the race of a potential renter and the chance of being shown section A.