Assignment 3

1. In Vancouver, British Columbia, the probability of rain during a winter day is 0.58, for a spring day is 0.38, for a summer day is 0.25, and for a fall day is 0.53. Each of these seasons lasts one quarter of the year.
a. What is the probability of rain on a randomly chosen day in Vancouver?
b. If you were told that on a particular day it was raining in Vancouver, what would be the probability that this day would be a winter day?

ANSWER
1. **(a)** Pr[rain on random day in Vancouver] = (0.25 × 0.58) + (0.25 × 0.38) + (0.25 × 0.25) + (0.25 × 0.53) = 0.435.
(b) Pr[winter| raining] = Pr[raining| winter] × Pr[winter] / Pr[raining] = 0.58 × 0.25 / 0.435 = 0.333

2. Three variants of the gene encoding the β-globin component of hemoglobin occur in the human population of the Kassena-Nankana district of Ghana, West Africa. The most frequent allele, A, occurs at frequency 0.83. The two other variants, S ("sickle cell") and C, occur at frequency 0.04 and 0.13, respectively (Ghansah et al. 2012). Each individual has two alleles, determining its genotype at the β-globin gene. Assume that knowing the identity of one of the alleles of any individual provides no information about the identity of the second allele (i.e., alleles occur independently in individuals).

a. CC individuals, having two copies of allele C, are slightly anemic. What is the probability that a randomly sampled individual from the population has two copies of the C allele (in other words, what is the probability that the individual's first allele is C and his/her second allele is also C)?
b. What is the probability that a randomly sampled individual is a homozygote (has two copies of the same allele)?
c. Compared with AA individuals, AS and AC individuals are largely resistant to malaria, which is endemic to the region. They also experience fewer deleterious side effects than SS and CC

individuals. What is the probability that a randomly sampled individual is AS? (Remember that if an individual can be AS by getting A from mom and S from dad or by getting S from mom and A from dad.)
d. What is the probability that a randomly sampled individual is AS or AC?

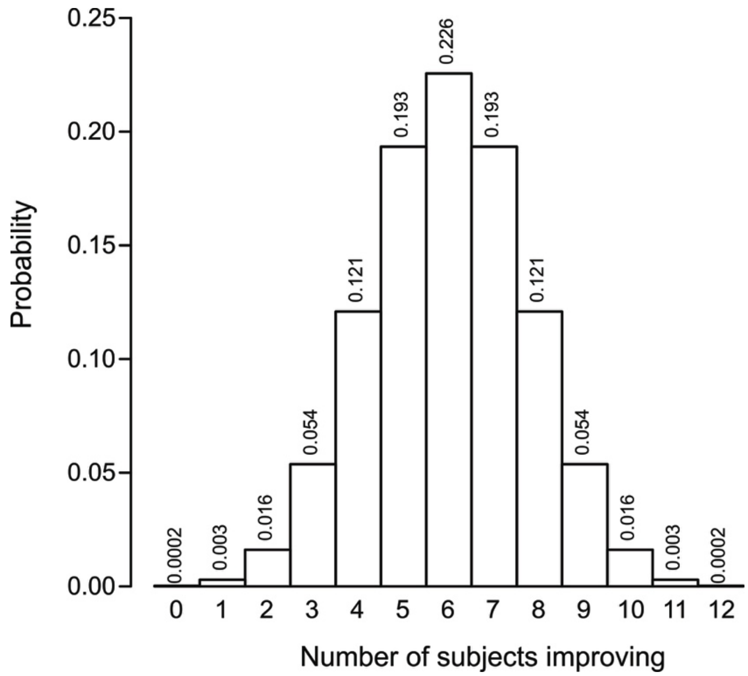
ANSWER
2. **(a)** Pr[*CC*] = Pr[*C*] Pr[*C*] = 0.13^2 = 0.0169.
(b) Pr[homozygote] = Pr[*AA*] + Pr[*SS*] + Pr[*CC*] = $0.83^2 + 0.043^2 + 0.13^2$ = 0.708.
(c) Pr[*AS*] = Pr[*A* then *S*] + Pr[*S* then *A*] = (0.83)(0.04) + (0.04)(0.83) = 0.0664.
(d) Pr[*AS* or *AC*] = Pr[*AS*] + Pr[*AC*] = 0.0664 + 2(0.83)(0.13) = 0.2822.

3. Identify whether each of the following statements is more appropriate as a null hypothesis or an alternative hypothesis:
a. Hypothesis: The number of hours that grade school children spend doing homework predicts their future success on standardized tests.
b. Hypothesis: King cheetahs on average run the same speed as standard spotted cheetahs.
c. Hypothesis: The mean length of African elephant tusks has changed over the last 100 years.
d. Hypothesis: The risk of facial clefts is equal for babies born to mothers who take folic acid supplements compared with those from mothers who do not.
e. Hypothesis: Caffeine intake during pregnancy affects mean birth weight.

ANSWER
3. **(a)** Alternative hypothesis
(b) Null hypothesis
(c) Alternative hypothesis
(d) Null hypothesis
(e) Alternative hypothesis

4. About 30% of people cannot detect any odor when they sniff the steroid androstenone, but they can become sensitive to its smell if exposed to the chemical repeatedly. Does this change in sensitivity happen in the nose or in the brain? Mainland et al. (2002) exposed one nostril of each of 12 non-detector subjects to androstenone for short periods every day for 21 days. The other nostril was plugged and had humidified air flow to prevent androstenone from entering. After the 21 days the researchers found that 10 of 12 subjects had improved androstenone-detection accuracy in the *plugged* nostril, whereas 2 had reduced accuracy. This suggested that increases in sensitivity to androstenone happen in the brain rather than the nostril, since the epithelia of the nostrils are not connected. The authors carried a statistical hypothesis test of whether accuracy in fact changed. Let p refer to the proportion of non-detectors in the population whose accuracy scores improve after 21 days. Under the null hypothesis, $p = 0.5$ (as many subjects should improve as deteriorate in their accuracy after 21 days). The alternative hypothesis is that $p \neq 0.5$ (the proportion of subjects increases or decreases after 21 days).
a. Did the authors carry out a one- or a two-sided test? What justification might they provide?

b. The accompanying figure shows the null distribution for the number of subjects out of 12 having an improved accuracy score. The probability of each outcome is given above the bars. To what do these probabilities refer?
c. What is the test statistic for the test?
d. What is the *P*-value for the test?
e. What is the appropriate conclusion? Use significance level α = 0.05.



ANSWER
4. **(a)** Two-sided test. The alternative hypothesis allows for the possibility that accuracy scores will decline as well as that they will increase. This is justified if an increase in accuracy is not the only feasible alternative outcome.
(b) Each number refers to the probability of obtaining exactly the number indicated of subjects improving *if the null hypothesis is true*.
(c) The number of subjects improving (the value 10 was observed).
(d) Pr[number $\geq$ 10] = 0.016 + 0.003 + 0.0002 = 0.0192. P = 2 (0.0192) = 0.038 (because it is a two-tailed test).
(e) Since *P* < 0.05, reject the null hypothesis. Accuracy scores improve in the plugged nostril after 21 days of training.

5. Refer to problem 4. The researchers also found that 11 of 12 subjects showed improved accuracy in the *exposed* nostril after 21 days.

a. Carry out the 4 steps of hypothesis testing (state hypotheses, compute test statistic, determine *P*-value, draw appropriate conclusion), to test whether accuracy scores changed after 21 days in the exposed nostrils of subjects. Use significance level $\alpha = 0.05$.
b. Why it would be useful also to provide a confidence interval for the proportion of subjects improving?

ANSWER
5. **(a)** 1) H_0: $p = 0.5$ (as many subjects improve as deteriorate). H_A: $p \neq 0.5$ (the proportion changes). 2) Test statistic is the number of subjects improving accuracy (11 is observed). 3) $P = 2(0.003 + 0.0002) = 0.006$. 4) Reject H_0, conclude accuracy scores improve.
(b) To estimate and put bounds on the most plausible values for the true proportion of subjects improving, *p*.

6. One classical experiment on ESP (extrasensory perception) tests for the ability of an individual to show telepathy—that is, to read the mind of another individual. This test uses five cards with different designs, all known to both participants. In a trial, the "sender" sees a randomly chosen card and concentrates on its design. The "receiver" attempts to guess the identity of the card. Each of the five cards is equally likely to be chosen, and only one card is the correct answer at any point.
a. Out of 10 trials, a receiver got four cards correct. What is her success rate? What is her expected rate of success, assuming she is only guessing?
b. Is her higher actual success rate reliable evidence that the receiver has telepathic abilities? Carry out the appropriate hypothesis test.
c. Assume another (extremely hypothetical) individual tried to guess the ESP cards 1000 times and was correct 350 of those times. This is very significantly different from the chance rate, yet the proportion of her successes is lower than the individual in part (a). Explain this apparent contradiction.

ANSWER
6. **(a)** 4 correct out of 10 = 0.4. If guessing, expected success rate = 1 /5 = 0.2
(b) Using the binomial test. H_0: The receiver had a probability of success of 1/5 ($p = 1/5$). H_A: The receiver's probability of success is not 1/5 ($p \neq 1/5$). We must calculate the probability of obtaining 4, 5, 6 . . . 10 correct results, and sum these together. For a two-tailed test, multiply the sum by 2. For four successes, $\Pr[4] = \binom{10}{4} 0.2^4 (1 - 0.2)^6 = 0.088$. Similar calculations produce the following table:

No. successes	**Probability**
4	0.0880804
5	0.0264241
6	0.0055050
7	0.0007864
8	0.0000737

9	0.0000041
10	0.0000001
Sum	0.1208739

For the two-tailed test, $P = 2(0.12) = 0.24$. Since $P > 0.05$, we do not reject the null hypothesis that the probability of success was 0.2.
(c) As the number of trials increases, the standard error of the estimated proportion declines (as the number of trials is in the denominator). In other words, the precision of the estimate of the proportion improves. With larger sample sizes, smaller differences can be detected.

7. Biff and Dilara were having an argument over what fraction of people would likely go out of their way to drive over a live organism if it were standing innocently by the side of the road. Dilara, whose heart is pure, guessed that fewer than 2% of people would behave that badly, roughly the proportion of people who score as psychopaths in standard testing. Biff, who isn't revealing what he knows, guessed that the fraction would be higher, perhaps 5%. To settle the debate they analyzed data from an experiment in which a rubber facsimile of either a turtle, a tarantula spider, a snake, or a leaf were placed on the paved shoulder of a two-way road[1] (Rober 2012). Of 1000 vehicles observed to drive by, 60 swerved onto the shoulder in an effort to drive over the rubber organism. Let's assume (perhaps unrealistically) that each vehicle represents an independent trial and that the probability of someone attempting to flatten the rubber organism is the same for each organism. Are these data consistent with a fraction of 2%? Are they consistent with a fraction of 5%?

R exercise
Use R to check the above answer.

ANSWER
7.$n = 1000$, $\hat{p} = 60/1000 = 0.06$, $p' = 0.0618$, the 95% confidence interval for p is $0.047 < p < 0.077$. The guess of 2% is outside the most-plausible range (too low), but 5% falls within the range and so is consistent with the data.

8. In North America, between 100 million and 1 billion birds die each year by crashing into windows on buildings, more than any other human-related cause. This figure represents up to 5% of all birds in the area. One possible solution is to construct windows angled downwards slightly, so that they reflect the ground rather than an image of the sky to a flying bird. An experiment by Klem et al. (2004) compared the number of birds that died as a result of vertical windows, windows angled 20 degrees off vertical, and windows angled 40 degrees off vertical. The angles were randomly assigned with equal probability to six windows and changed daily; assume for this exercise that windows and window locations were identical in every respect except angle. Over the course of the experiment, 30 birds were killed by windows in the vertical

orientation, 15 were killed by windows set at 20 degrees off vertical, and eight were killed by windows set at 40 degrees off vertical.

a. Clearly state an appropriate null hypothesis and an alternative hypothesis.

b. What proportion of deaths occurred while the windows were set at a vertical orientation?

c. What statistical test would you use to test the null hypothesis?

d. Carry out the statistical test from part (c). Is there evidence that window angle affects the mortality rates of birds?

R exercise

e. Load the data file. Calculate the proportion of deaths for the three differences angles.

f. make a data frame with in the first column the observed deaths, the second the expected deaths and in the third column $\frac{(Observed - Expected)^2}{Expected}$ (recreate a χ^2 table) and calculate the χ^2 value. Check if this is the same as you calculated manually.

g. R has a build in function as well to do the goodness of fit test *chisq.test*(). Look at the help that comes with it and try to figure out how it works. HINT: you only need the observed data if you assume equal probability of occurrence for each category.

ANSWER

8. **(a)** H_0: Windows kill the same number of birds per time period at any angle.

H_A: Windows angled towards the ground kill a different number of birds per time period than windows at the vertical.

(b) 30 / 53 were killed by windows at the vertical, or 0.566.

(c) We can use a goodness of fit test for the null hypothesis.

(d) The null hypothesis implies windows at each angle should kill 33% of the birds.

window angle	observed deaths	expected deaths	$\frac{(Observed - Expected)^2}{Expected}$
(vertical) 0	30	17.67	8.6
20	15	17.67	0.4
40	8	17.67	5.3
Sum	53	53.01	χ^2 = 14.3

We had three categories, no estimated parameters, so df= 2. χ^2 =14.3 > 13.92, the critical value for α = 0.001, so P < 0.001 and we reject H_0. Window angle does influence bird mortality (P = 0.0008).

9. In snapdragons, variation in flower color is determined by a single gene (Hartl and Jones 2005). *RR* individuals are red, *Rr* (heterozygous) individuals are pink, and *rr* individuals are white. In a cross between heterozygous individuals, the expected ratio of red-flowered: pink-flowered: white-flowered offspring is 1:2:1.
a. The results of such a cross were 10 red-, 21 pink-, and nine white-flowered offspring. Do these results differ significantly (at a 5% level) from the expected frequencies?
b. In another, larger experiment, you count 100 times as many flowers as in the experiment in part (a) and get 1000 red, 2100 pink, and 900 white. Do these results differ significantly from the expected 1:2:1 ratio?
c. Do the proportions observed in the two experiments [i.e., in parts (a) and (b)] differ? Did the results of the two hypothesis tests differ? Why or why not?

R exercise
d. Redo the analyses of a and b above using the *chisq.test*() but note that the predicted frequency is not equal for all categories. First try yourself to figure out how you can provide the function with a list of expected frequencies (which have to be probabilities). A hint it given at the end of the next question if you can't find it.
e. Manually look for the sample size in which the difference found is significant (don't spend ages on this, just roughly). Can you think what this result means in the light of people 'accepting a null-hypothesis' ? (Remember you never accept a H_0.)

ANSWER
9. **(a)** H_0: The frequency distribution of genotypes follows a 1:2:1 ratio. H_A: The frequency distribution of genotypes does not follow a 1:2:1 ratio.

	Observed Frequency	Expected Frequency	$\frac{(Observed - Expected)^2}{Expected}$
red	10	10	0.00
pink	21	20	0.05
white	9	10	0.10

$\chi^2 = 0.15$

For 2 *df*, the critical value is 5.99, which is larger than $\chi^2 = 0.15$, so we do not reject H_0 (P = 0.93). The data are consistent with a 1:2:1 ratio.

(b)

	Observed Frequency	**Expected Frequency**	$\frac{(Observed - Expected)^2}{Expected}$
red	1000	1000	0
pink	2100	2000	5
white	900	1000	10
Sum	4000	4000	$\chi^2 = 15$

Yes, now we see that χ^2 is greater than the critical value, so we reject H_0 (P = 0.00055).

(c) The proportions do not differ, but the tests of the two hypotheses differ. Even small proportional differences can be statistically significant with sufficiently large sample size. This is why we can never “accept” the null hypothesis.