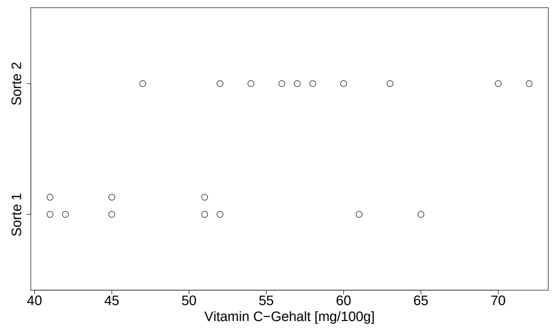
Dirk Metzler SS 2026

STATISTICS FOR EES AND MEME — EXERCISE SHEET 5

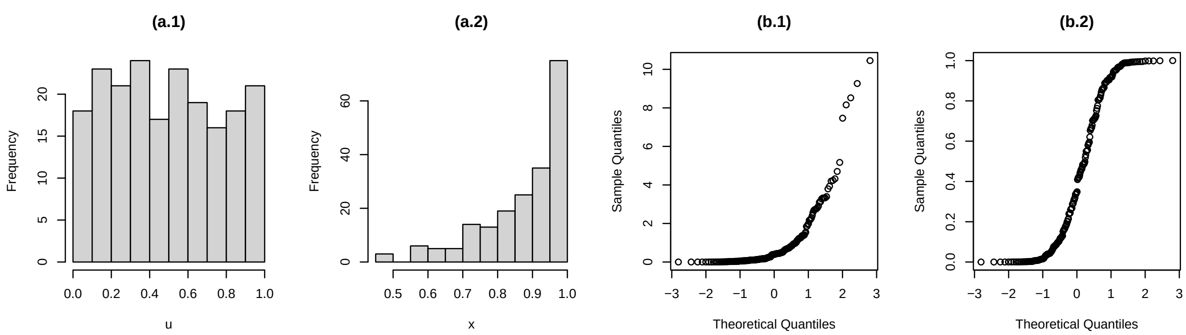
1. The vitamin C contents of two types of cabbage were surveyed in 10 samples per type:



For the samples of type 1 a mean content of $\mu_1 = 49.4$ mg per 100g with standard deviation $s_1 = 8.33$ was found, for type 2 the values were $\mu_2 = 58.9$ and $s_2 = 7.74$. Test the hypothesis that the mean vitamin C contents is the same for both types with a t-test. You may assume that the true variances are equal.

2.

(a) Sketch normal-quantile–quantile ("qqnorm") plots for the distributions shown in histograms a.1 and a.2.



(b) Sketch density plots of the distributions for which qqnorm plots are shown in figures b.1 and b.2.

(c) In R you can use the command `rnorm(n)` to simulate a sample of size n from the standard normal distribution and the command `rgamma(n, shape=s)` to simulate a sample from the so-called gamma distribution with shape parameter s. The R command `dgamma(x, shape=s)` returns the density of the gamma distribution at x. Plot this density for the shape parameter values 0.2, 1, 2, 3 and 10. Then simulate samples of various sizes n and, in the case of the gamma distribution, with $s = 2$ and and $s = 3$ and explore how large n must be to make the qqnorm plots of the respective gamma distribution clearly distinguishable from those of samples from the normal distribution.

3. The t-test requires that the sample consists of independent observations. Explore how robust the t-test is against violations of this requirement. To this end, simulate data `x` by

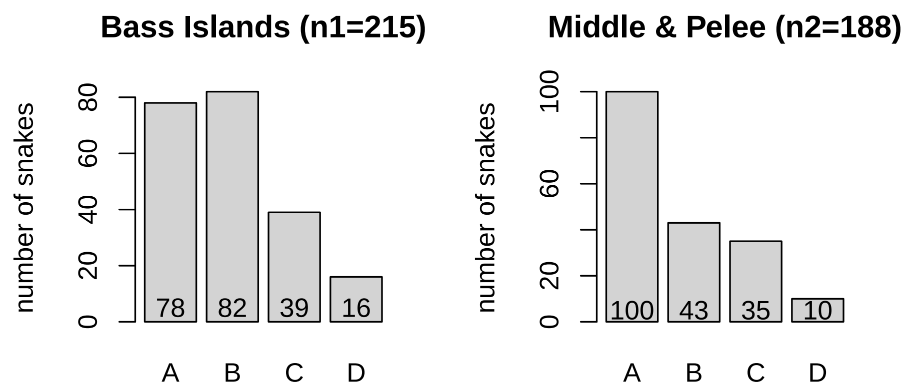
```
x <- rnorm(20) + sample(rnorm(10),20,replace=TRUE).
```

(a) Explain what this command does.

(b) What is the expected value of x?

(c) Perform t-tests on x for the null hypothesis that it is sampled from a normal distribution with mean 0. Repeat this with 10000 different simulated data sets x. How often is the null hypothesis rejected?

4. The markings of the water snake *Nerodia sipedon* in Lake Erie can be grouped into four classes. On the mainland all snakes have strong markings (D). On the islands many snakes have no markings (A) or only weak marking of two different types (B and C). Frequencies of the four markings in samples from the Bass islands and from Middle Island and Pelee Island are as follows.



(a) Let H_0 be the hypothesis that the distribution of marking classes were the same for all islands. What is the expected number of snakes of class A sampled from the Bass Islands if we assume that H_0 is true?

(b) We apply a Chi-square test with null hypothesis H_o. How much does the table entry (Bass Islands : class A) contribute to the test statistic X^2?

(c) If we sum $(O_i - E_i)^2/E_i$ for all table entries we find $X^2 = 14.74$. How many entries does the table have? How many degrees of freedom does X^2 have?

(d) What is the value of the quantile to which you have to compare the p value to check for significance on the 5% level?

(e) Which sentence is appropriate to summarize the result of the test? (One and only one is correct.)

   (a) The frequencies of the four markings differ significantly between the two populations.

   (b) The frequencies of the markings do not differ between the two populations.

   (c) We cannot reject the null hypothesis that the frequencies of the markings are the same in the two populations.

   (d) We reject the null hypothesis that the frequencies of the markings differ between the two populations.

   (e) The frequencies of the markings differ between the two populations.

5. (Inspired by Ouattara et al., 2009, Animal Behaviour 78:35–44) Females of Campbell's mona monkeys can produce several different alarm calles, the main types being "wak-oos", "hoks" and "trill" calls. In a research project over several months you present three different visual predator models (eagle, leopard and snake) several times per day in randomized orderto a group of monkeys, and you record for two females (A and B) how often which alarm call was used in which case.

(a) In 87 cases female A was the one who gave the warning call. The following table shows how often she used the different calles in the different cases:

	wak-oos	hoks	trill
eagle	16	5	7
leopard	6	8	20
snake	7	15	3

Perform a statistical test to check whether A's choice of the warning call depends on the species of the predator model. Summarize the result of the test in a sentence that refers explicitly to the application context.

(b) How much did the case eagle/hoks contribute to the test statistic?

(c) Explain the number of degrees of freedom and find in the quantile table the threshold value for significance on the 5% level.

(d) (Advanced!) In 72 cases female B was the one who called, and the the following table shows how often she used which call in which case:

	wak-oos	hoks	trill
eagle	8	4	7
leopard	5	11	15
snake	4	10	8

Test for differences between females A and B in how the distribution of call types depends on the predator species. Summarize the result of the test in a sentence that refers explicitly to the application context.

6. In a study[1] about the alkaline-phosphatase gene, three distinguishable alleles "S", "I" and "F" were found. The following numbers of genotypes were observed in 332 persons: SS: 141, SF: 111, FF: 28, SI: 32, FI: 15, II: 5.

(a) Compute the relative allele frequencies for S, I and F.

(b) Use the allele frequencies to compute the expected values for the genotypes in a sample of 332 persons, assuming a Hardy-Weinberg equilibrium for this gene.

(c) Is the observed deviation from Hardy-Weinberg equilibrium significant?

7. To test whether a pharmaceutical increases the reaction time, the latter was tested with nine test persons. Five of them were randomly selected and took the drug. Their reaction times in seconds were 0.78, 0.66, 0.86, 0.90, 0.83. The other four persons had the reaction times 0.82, 0.62, 0.63, 0.69. Apply the Wilcoxon rank sum test and an appropriate t-test by hand (that is, without using the R commands `wilcox.test` and `t.test`) to test whether the drug increases the reaction time.

[1] Harris (1966) Enzyme polymorphism in Man. *Proc. Roy. Soc. B* **164**:1153-64

8. (Optional; for those who watched the videos on the two-sample Wilcoxon test) Test the power (that is the ability to reject the null hypothesis if it is wrong) and robustness (against violations of the requirements on the distribution) of the two-sample t-test and the Wilcoxon rank sum test:

(a) Generate two normally distributed samples of size n, one of them with true mean 0 and one with true mean μ (with R: `rnorm(`n`,mean=`μ`)`). How does the probability that the two-sample t-test or the Wilcoxon rank sum test rejects the null hypothesis "the population means are equal" with significance level $\alpha = 0.05$ on the values of n and μ? Explore this for $\mu \in \{0, 0.5, 2\}$ and $n \in \{5, 10, 20\}$ by repeated simulations.

(b) Repeat (a) with a slight modification. Generate one sample with `rnorm(`n`,mean=0)` again, but the other one with `rnorm(`n`,mean=`μ`, sd=`σ`)`, where σ takes a value between 2 and 10.

(c) In another series of simulations generate one sample with `rexp(`n`, rate=1)` and the other one with `rexp(`n`, rate=`r`)`. For r try the values 1, 0.5, and 0.1.