Exercise H9.1. *(Exercise 8.59 e, p. 399 textbook).* A famous medical experiment was conducted by Joseph Lister in the late 1800s. Mortality associated with surgery was quite high and Lister conjectured that the use of a disinfectant, carbolic acid, would help. Over a period of several years Lister performed 75 amputations with an without using carbolic acid. The data are

		Carbolic acid	*used ?*
		Yes	No
Patient	Yes	34	19
lived ?	No	6	16

Use these data to test whether the use of carbolic acid is associated with patient mortality.

Solution: By CLT, under $H : p_1 = p_2 = p$, we have

$$T = \frac{\left(\hat{p_1} - \hat{p_2}\right)^2}{p\,(1-p)\left(\frac{1}{n_1} + \frac{1}{n_2}\right)} \overset{app}{\sim} \chi_1^2.$$

And $\hat{p_1} = 34/40$, $\hat{p_2} = 19/35$, $\hat{p} = (34+19)/(40+35) = 53/75$. Then $T = 8.495$. Since $\chi^2_{1,0.05} = 3.84$, we can reject H at $\alpha = 0.05$.

Exercise H9.2. *(Adapted from exercise 8.60, p. 399 textbook)* Let $\mathbf{Z} = (Z_1, \ldots, Z_k)$ have a multinomial law $\mathfrak{M}_k(n, \mathbf{p})$ with unknown $\mathbf{p} = (p_1, p_2, \ldots, p_k)$, where $k > 2$. Consider hypotheses on the first two components
$H : p_1 = p_2$
$K : p_1 \neq p_2$

A test that if often used, called *McNemar's Test*, rejects H if

$$\frac{(X_1 - X_2)^2}{X_1 + X_2} > \chi^2_{1;1-\alpha} \tag{1}$$

where $\chi^2_{1;1-\alpha}$ is the lower $1-\alpha$ quantile of the distribution $\chi^2_{1;1-\alpha}$[1].

(i) Find the maximum likelihood estimator $\hat{\mathbf{p}}$ of the parameter $\mathbf{p}$ under the hypothesis.

Hint: Proposition 8.1, p. 84 handout gives the MLE under no restriction on $\mathbf{p}$, except being a probability vector; $\hat{\mathbf{p}}$ was written $\hat{\mathbf{p}} = (\hat{p}_1, \ldots, \hat{p}_{k-1}, 1 - \sum_{j=1}^{k-1} \hat{p}_j)$ and the likelihood was maximized in $.\hat{p}_1, \ldots, \hat{p}_{k-1}$. Under the additional restriction $p_1 = p_2$, write $\hat{\mathbf{p}} = (\hat{p}_*, \hat{p}_*, \hat{p}_3, \ldots, \hat{p}_{k-1}, 1 - \sum_{j=1}^{k-1} \hat{p}_j)$ and maximize in $\hat{p}_*, \hat{p}_3, \ldots, \hat{p}_{k-1}$.

[1]The textbook writes an upper α-quantile, called $\chi^2_{1,\alpha}$ there; it coincides the lower quantile $\chi^2_{1;1-\alpha}$.

(ii) Show that the appropriate χ^2-statistic with estimated parameter $\hat{\mathbf{p}}$ (maximum likelihood estimator under H as above), as defined in relation (8.27), p.100 handout, coincides with McNemar's statistic (1) (exact equality, not approximate with an error term).

Comment: It follows that McNemar's test is the χ^2-test for this problem and is an asymptotic α-*t*est, cf. Theorem 8.12, p.100 handout.

Solution: **i**) Under $H : p_1 = p_2 = p$, we have

$$\log L(\mathbf{p}|\mathbf{x}) = x_1 \log p + x_2 \log p + x_3 \log p_3 + ... + x_n \log \left(1 - 2p - \sum_{i=3}^{n-1} p_i\right).$$

Taking logs and differentiating yield the folowing equations for the MLEs:

$$\frac{\partial \log L}{\partial p} = 2(\frac{x_1 + x_2}{2p} - \frac{x_n}{\left(1 - 2p - \sum_{i=3}^{n-1} p_i\right)}) = 0,$$

$$\frac{\partial \log L}{\partial p_i} = \frac{x_i}{p_i} - \frac{x_n}{\left(1 - 2p - \sum_{i=3}^{n-1} p_i\right)} = 0, i = 3, ..., n-1.$$

This implies

$$\frac{x_1 + x_2}{2p} = \frac{x_3}{p_3} = ... = \frac{x_n}{\left(1 - 2p - \sum_{i=3}^{n-1} p_i\right)} = \frac{m}{1},$$

where $m = \sum_{i=1}^{n} x_i$, and $1 = 2p + p_3 + ... + \left(1 - 2p - \sum_{i=3}^{n-1} p_i\right)$.

Thus $\hat{p} = \dfrac{x_1 + x_2}{m}, \hat{p_i} = \dfrac{x_i}{m}, i = 3, .., n-1.$

ii) Except for the first and second cells, we have expected=observed, since both are equal to x_i.

Thus we get

$$\begin{aligned} \sum \frac{(observed - \exp ected)^2}{\exp ected} &= \frac{\left(x_1 - \frac{x_1+x_2}{2}\right)^2}{\frac{x_1+x_2}{2}} + \frac{\left(x_2 - \frac{x_1+x_2}{2}\right)^2}{\frac{x_1+x_2}{2}} \\ &= \frac{(x_1 - x_2)^2}{x_1 + x_2}. \end{aligned}$$