Solution Sheet 8

1. **a)** Showing a breach of the limit is more difficult if the standard deviation is estimated from data. The test statistic

$$T = \frac{\overline{X} - 200}{\widehat{\sigma}/\sqrt{n}}$$

follows a t distribution with $n - 1$ degrees of freedom, and in particular, this distribution is wider than that of

$$Z = \frac{\overline{X} - 200}{\sigma/\sqrt{n}}$$

with σ known. Thus the critical value of the test statistic is larger than when σ is known, and the t test thereby has less power than the z-test.

b) Formally, the t-test is carried out as follows:

Model assumption:	X_i: $i-$th measurement of ammonia levels. X_i i.i.d. $\mathcal{N}(\mu, \sigma^2)$ where σ is unknown.
Null hypothesis H_0:	X_i i.i.d. $\mathcal{N}(\mu_0, \sigma^2)$ with $\mu_0 = 200$
Alternative hypothesis H_A:	X_i i.i.d. $\mathcal{N}(\mu, \sigma^2)$ with $\mu > 200$ (one-sided)
Rejection set:	From the table: $\mathcal{K} = \{t : t_{15,0.95} > 0.95\} =]1.753, \infty]$. This corresponds to a rejection set $]204.38, \infty]$ for $\overline{X}$.
Value of the test statistic:	$t = \frac{\overline{x}-\mu_0}{\widehat{\sigma}/\sqrt{n}} = \frac{204.2-200}{\widehat{\sigma}/\sqrt{16}} = 1.68$
Outcome:	$1.68 \notin \mathcal{K}$; thus the null hypothesis cannot be rejected using th The breach of the limit has not been statistically proven.

Although the discrepancy between the t-test and the z-test is small, it leads to the null hypothesis not being rejected in this example.

c) The data might not be normally distributed. Non-normal data give the t-test bad power. In such cases a sign or Wilcoxon test would be more suitable.

2. **a)** Let X denote the lead content of lettuces. We have:

$$X \sim \mathcal{N}(\mu, \sigma^2), \text{ where } \sigma = 10 \text{ and } \mu \text{ is unknown.}$$

The average $\overline{X}$ of $n = 10$ samples also follows a normal distribution, with standard deviation $\sigma/\sqrt{n}$,

$$\overline{X} \sim \mathcal{N}(\mu, \sigma^2/n).$$

As $\Phi(2.58) = 0.995$ (given in the hint), 99% of all observations of the standardized random variable

$$\frac{\overline{X} - \mu}{\sigma/\sqrt{n}}$$

lie in the interval $[-2.58, 2.58]$. Thus 99% of all observations of $\overline{X} - \mu$ lie in the interval $[-2.58 \cdot \sigma/\sqrt{n}, 2.58 \cdot \sigma/\sqrt{n}]$. This this that a 99% confidence interval for μ is given by

$$\left[\widehat{\overline{X}} - 2.58 \cdot \frac{\sigma}{\sqrt{n}}, \widehat{\overline{X}} + 2.58 \cdot \frac{\sigma}{\sqrt{n}}\right],$$

where $\widehat{\overline{X}}$ is the observed average, which here is $\widehat{\overline{X}} = 31$. Using the known values $\sigma = 10$ and $n = 10$, we obtain the 99% confidence interval $[22.84, 39.16]$

b) By Part a) we see that the width of the confidence interval decreases as $1/\sqrt{n}$ does when the sample size n is raised. Thus in order to halve the width of the confidence interval, we need four times as many observations, i.e. $4 \cdot 10 = 40$.
The width of the 99% confidence interval is

$$2 \cdot 2.58 \cdot \frac{\sigma}{\sqrt{n}}.$$

To reduce this width to 1 ppb, we must raise the sample size n accordingly:

$$\begin{aligned} 2 \cdot 2.58 \cdot \frac{\sigma}{\sqrt{n}} &\leq 1 \\ \rightarrow \qquad 51.6 &\leq \sqrt{n} \\ \rightarrow \qquad n &\geq 2663. \end{aligned}$$

We require at least 2663 observations, if we are to obtain a 99% confidence interval whose width does not exceed 1 ppb.

c) The standardized random variable

$$\frac{\overline{X} - \mu}{\widehat{\sigma}/\sqrt{n}}$$

, which uses the estimate $\widehat{\sigma}$ of σ, no longer follows a normal distribution (unlike when σ is fixed and known). Instead, it follows a t distribution with 9 degrees of freedom. The 99.5% quantile of this distribution is 3.25 (see the table of the t distribution). Thus 99% of all observations of

$$\frac{\overline{X} - \mu}{\widehat{\sigma}/\sqrt{n}}$$

lie in the interval $[-3.25, 3.25]$, and we obtain the confidence interval

$$\left[\widehat{\overline{X}} - 3.25 \cdot \frac{\widehat{\sigma}}{\sqrt{n}}, \widehat{\overline{X}} + 3.25 \cdot \frac{\widehat{\sigma}}{\sqrt{n}}\right].$$

For $\sigma = 10$ and $n = 10$, we obtain the confidence interval $[20.72, 41.28]$. Comparing this with Part a), we find that the confidence interval has been scaled by a factor 3.25/2.58, i.e. by roughly 0.26%.

3. The vibrations caused by construction work may cause cracks in buildings to widen, but they are unlikely to make them narrower. Thus if we equate "widening cracks" with "damage to buildings", we can carry out a one-sided test.

D_i = After - Before

Null hypothesis H_0:	$D_i \sim \mathcal{F}_o$ with median $\mu_0 = 0$, independently for i=1,2,...,n (No change)
Alternative hypothesis H_A:	$D_i \sim \mathcal{F}$ with median $\mu > 0$, independently for i=1,2,...,n
Test statistic:	U = Anzahl $\{i \mid D_i > 0\}$
Under H_0, we have:	$U \sim \mathcal{B}(12, 0.5)$
Rejection set:	$\mathcal{K} = \{U \geq 10\}$

n	12	11	10	9
p_n	0.00024	0.0029	0.0161	0.0537
$\sum_{i=n}^{12} p_i$	0.00024	0.0032	0.0193	0.073

We have $u = 9 \notin \mathcal{K}$. The null hypothesis is **not rejected**, i.e. the changes during the building work are not statistically significant.