Lesson 1

Ex. A a) $P(X < 3.5) = 0.16$ b) $P(\overline{X} < 3.5) = 0.02$.

Ex. B **We have pairwise measurements. We construct differences $d_i = x_i - y_i$.**
Model: The r.v. $D_i \sim N(\mu_D, \sigma)$, where μ_D describes the systematic difference between the models.
$I_{\mu_D} = (\bar{d} \pm 2.26 \cdot s_d/\sqrt{10}) = (-0.11; 5.11)$, where s_d is sample standard deviation fpr d_i-values. Since $0 \in I_{\mu_D}$, we cannot conclude that there is a systematic difference between the methods.

Ex. C **a) Test statistics $\frac{\bar{x}-5}{0.87/\sqrt{24}} = -2.20 > -2.33$ H_0 cannot be rejected.**
b) Power $h(4.5) = \Phi(0.486) \approx 0.69$.
c) Test in (a) is better, since when we reject in (a), we can drink the water with 99% sure.

Ex. D a) $I_{\mu_1-\mu_2} = (\bar{u} - \bar{v} - 1.72s\sqrt{\frac{22}{117}}, \infty) = (0.864, \infty)$
Allergic people on average have higher values than non-allergic.
b) $P(X_i > 50) \approx 0.79$ and $P(Y_j > 50) \approx 0.24$.

Ex. E a) $\mu > 2.165$
b) $I_\mu = (2.29, \infty)$; condition in a) is with high probability satisfied.
c) Fact: $15S^2/\sigma^2 \sim \chi^2(15)$ and it gives $I_\sigma = (0.392, 0.882)$, so $\sigma = 0.5$ seems to be reasonable assumption for our model.

Ex. F $I_\sigma = (0, s\sqrt{21/11.59}) = (0, 0.317)$

Ex. G **Difference $\mu_1 - \mu_2$ describes the systematic difference between the indicators**

$$I_{\mu_1-\mu_2} = (\bar{x} - \bar{y} \pm 2.02 \cdot s \cdot \sqrt{\frac{1}{16} + \frac{1}{26}}) = (-0.00024; 0.00114).$$

We see that $0 \in I_{\mu_1-\mu_2}$ and that the interval is short. The systematic difference seems negligible.

Ex. H **a) Test statistic $\frac{\bar{x}-2.5}{0.32/\sqrt{15}} = -0.97 > -1.645$; H_0 can not be rejected.**
b) $1 - \Phi(0.44) \approx 0.33$, i.e. poor power for $\mu = 2.40$.

Ex. I **a) The observed points follow the curved curve much better. The straight line in the first plot seems to be systematically wrong in relation to the observed values.**
b) $I_{\beta_2} = (\hat{\beta}_2 \pm t \cdot s \cdot \sqrt{h_{22}}) = (-7.11; -4.10)$. We see that $0 \notin I_{\beta_2}$. Hence, x^2 is useful as an explanatory variable.
c) For the estimated regression relationship $x = 10.22$ is the value that gives highest reduction of the phosphate. This is only an estimate of the optimum x-value.
d) $\hat{m}_{10} - \hat{m}_{11} = -\hat{\beta}_1 - 21\hat{\beta}_2 = 3.188$. Hence, pH=10 seems to be better than pH=11.