December 21, 2018

# EXAM IN HMET4100, 6 DECEMBER 2018: SUGGESTED SOLUTIONS AND COMMENTS 

Knut R. Wangen<br>k.r.wangen@medisin.uio.no

## Introduction

The exam was arranged digitally on the Inspera platform and consisted of 9 questions. Three of the questions were multiple choice questions where the candidates should choose one out of a set of alternative answers. The remaining questions were in "essay" format and the answers had to be written as text.

The multiple choice questions were assessed automatically, while the "essay" questions were assessed manually. The maximum score for each question was 10 points and the minimum was 0 points. In total, the maximum score was 90 points.

The following relation between grades (A-F) and points [0-90] was used: A: [81-90], B: [71-80], C: [55-70], D: [41-54], E: [30-40], F: [0-29].

The number of candidates was 52 and the grade distribution was as follows:

| Grade | No. of candidates | Candidates in \% |
| :---: | :---: | :---: |
| A | 9 | $17.3 \%$ |
| B | 8 | $15.4 \%$ |
| C | 12 | $23.1 \%$ |
| D | 10 | $19.2 \%$ |
| E | 7 | $13.5 \%$ |
| F | 6 | $11.5 \%$ |

Explanation of grades. After the grades are published, the candidates will have access to their individual scores by logging on to Inspera. The present suggested solution, which will also be available on Inspera, explains the general principles for how scores and grades were awarded. A candidate's detailed scores on the exam questions and the general suggested solution are together considered as the formal explanation of the grade, and no further formal explanations will be given.

Appeal against grades. Appeals against grades must be submitted following the procedure described on the course webpage (that is, not submitted to the course coordinator).

Informal feedback. Candidates who wish to receive additional feedback can ask the course coordinator for an informal talk (phone call or personal meeting). The candidates must then email the course coordinator 2-3 suggested times. These suggested times should be within normal office hours (0900h-1700h), with at least 48 hours notice, and within 11 January 2019.

## Suggested solutions

## Question 1

$P(X>4)=1-P(\leq 4)=1-0.9162 \approx \underline{\underline{0.084}}$.
Question 2

$$
P(X>30)=P\left(Z>\frac{30-37}{12}\right)=1-P(Z<-0.58)=\underline{\underline{0.719}} .
$$

## Question 3

a. The null hypothesis $\left(H_{0}\right)$ is that the mean values are the same for all three groups. The test statistic is

$$
F_{o b s}=\frac{M S G}{M S W}=\frac{103.78818}{76.3068247}=1.360
$$

The critical value $F_{K-1, n-K, \alpha}=F_{2,54, \alpha}$ is not tabulated but must between $F_{2,40, \alpha}$ and $F_{2,60, \alpha}$. For $\alpha=5 \%$ we have

$$
F_{2,40,0.05}=3.232 \quad \text { and } \quad F_{2,60,0.05}=3.150 .
$$

Thus, we do not reject $H_{0}$ at the $5 \%$ level, because $F_{\text {obs }}=1.36$ must be less than the critical value. The test result implies that we can continue to believe that the mean values are the same.
b. The null hypothesis is that the median values (or the distributions) are the same for the three patient groups.
The p -value of $0.0041=0.41 \%$ imply that we should reject the null hypothesis in the Kruaskal-Wallis test, even at the $1 \%$ significance level.
The one-way ANOVA and the Kruskal-Wallis tests result in different conclusions. One possible explanation is that the data are not normally distributed. This would violate the ANOVA test's assumptions, and imply that the ANOVA test results are not reliable.

## Question 4

The observed test statistic

$$
|t|=\frac{\bar{d}}{S . E \cdot(\bar{d})}=\frac{1320.455}{110.5779} \approx 11.94,
$$

while the critical value at the $1 \%$ significance level is $t_{10,0.005}=2.576$. Thus, the correct answer is: 11.94 , Reject the null hypothesis.

## Question 5

The main assumption we can consider is that the differences are drawn from a normally distributed population. ${ }^{1}$ If the population of differences is normally distributed and the sample size is large, we would expect the sample distribution of differences to be symmetric and bell-shaped.

The box plot on the right-hand side indicates that the distribution is symmetric because the median is fairly close to the midpoint of the box and the whiskers are of approximately equal length. The histogram with the normal density plot shows a concentration of observations in proximity of the sample mean. Neither graphs show any outliers.

The sample size $(n=11)$ is rather small, so we cannot expect the sample distribution to closely resemble the population distribution. However, based on the graphs at hand, it does not seem unreasonable to assume a normally distributed population. ${ }^{2}$

If the the population is not normally distributed, we could either use a sign test or a Wilcoxon signed rank test. Both are non-parametric tests and do not rely on the normality assumption. In the current application it would perhaps be most appropriate to use the Wilcoxon signed rank test, because we have a continuous numerical variable and because the numerical size of the differences seems relevant.

## Question 6

1. Assigning the value $\widehat{p}=0.5$, we obtain

$$
n=\frac{\widehat{p}(1-\widehat{p})\left(Z_{\alpha / 2}\right)^{2}}{\mathrm{ME}^{2}}=\frac{0.25 \cdot 1.96^{2}}{0.02^{2}}=\underline{\underline{2401}} .
$$

2. Let $p=0.15$ denote the assumed population proportion. If we let this value replace $\widehat{p}$ in the formula for the confidence interval we find

$$
n=\frac{p(1-p)\left(Z_{\alpha / 2}\right)^{2}}{\mathrm{ME}^{2}}=\frac{0.15 \cdot 0.85 \cdot 1.96^{2}}{0.02^{2}}=216.09
$$

that is, $\underline{\underline{217} \text { when rounding up. }}$

## Question 7

The test statistic has a standard normal distribution, and the critical value is $Z_{c}=1.645$. Assuming $H_{0}$ is true, the corresponding critical value for the sample mean is

$$
\bar{X}_{c}=Z_{c} \cdot \frac{\sigma}{\sqrt{n}}+\mu=1.645 \cdot \frac{2}{\sqrt{100}}+1=1.329 .
$$

[^0]The probability of type II error is

$$
\begin{aligned}
P\left(\bar{X}<\bar{X}_{c} \mid \mu=1.1\right) & =P\left(\frac{\bar{X}-1.1}{\sigma / \sqrt{n}}<\frac{\bar{X}_{c}-1.1}{\sigma / \sqrt{n}}\right) \\
& =P\left(Z<\frac{1.329-1.1}{2 / \sqrt{100}}\right)=P(Z<1.15)=\underline{\underline{0.8749}} .
\end{aligned}
$$

## Question 8

The Stata output states that the test was a "Two-sample $t$ test with equal variances".
We follow the notation from the output and let the numbers 1 and 2 denote the two hospitals, so that $\mu_{1}$ and $\mu_{2}$ denote their respective mean costs.
The Stata procedure tests the hypothesis $H_{0}: \mu_{1}-\mu_{2}=0$, that is, the difference in mean costs are hypothesized to equal zero, and we choose a two-sided alternative hypothesis, $H_{1}: \mu_{1}-\mu_{2} \neq 0$.
With $\alpha=0.05$ and $n_{1}+n_{2}-2=21+21-2=40$ degrees of freedom, we will reject $H_{0}$ if the absolute value of the test statistic is greater than the critical value $t_{40,0.025}=2.021$.

Three alternative solutions are considered sufficient:
Alternative 1. Performing the test, we calculate the pooled variance

$$
\begin{aligned}
& s_{p}^{2}=\frac{\left(n_{1}-1\right) s_{1}^{2}+\left(n_{2}-1\right) s_{2}^{2}}{n_{1}+n_{2}-2} \\
&=\frac{(21-1) 1.146603^{2}+(21-1) 0.6685858^{2}}{21+21-2}=0.880853
\end{aligned}
$$

and the test statistic

$$
T_{\text {obs }}=\frac{\bar{x}_{1}-\bar{x}_{2}}{\sqrt{\frac{s_{p}^{2}}{n_{1}}+\frac{s_{p}^{2}}{n_{2}}}}=\frac{6.956545-6.198018}{\sqrt{\frac{0.880853}{21}+\frac{0.880853}{21}}}=2.619 .
$$

We have that $\left|T_{\text {obs }}\right|=2.619>2.021$, thus we reject $H_{0}$ and conclude that the observed difference in means is statistically significant.

Alternative 2. The test statistic can be derived from the point estimate of the difference in means and the corresponding standard error, both reported in the "diff" line of the Stata output:

$$
T_{\text {obs }}=\frac{\text { point estimate }}{\text { standard error }}=\frac{0.758527}{0.2896389}=2.619 .
$$

The conclusion is similar to that in Alternative 1.
Alternative 3. The $95 \%$ confidence interval reported in the "diff" line of the Stata output is $(1.731449,1.343909)$. This interval does not include zero, which implies that the null hypothesis will be rejected when tested
against a two-sided alternative hypothesis at the $5 \%$ significance level (both hypothesis are as stated above). ${ }^{3}$

## Question 9

Three alternative solutions are suggested below. The first alternative is based on Stata's calculations of the expected value of the signed rank and the corresponding (adjusted) variance. However, Stata's calculations are based on a somewhat different method than the more straightforward method prescribed in the textbook, and the link between the two methods has not been explicitly treated in the organized teaching sessions. The two other alternatives are therefore also considered satisfactory.

Alternative 1. The observed $Z$-statistic is

$$
Z_{o b s}=\frac{T-E(T)}{\sqrt{\operatorname{Var}(T)}}=\frac{371.5-583}{9229.88}=-2.20
$$

Choosing a two-sided alternative hypothesis, the corresponding p-value is $2 P(Z<-2.20)=0.0278$. Thus, we reject the null hypothesis at the $5 \%$ significance level.

## Alternative 2.

$$
\begin{aligned}
E(T) & =\frac{n(n+1)}{4}=\frac{48(48+1)}{4}=588 \\
\operatorname{Var}(T) & =\frac{n(n+1)(2 n+1)}{24}=\frac{48(48+1)(2 \cdot 48+1)}{24}=9506 \\
Z_{o b s} & =\frac{T-E(T)}{\sqrt{\operatorname{Var}(T)}}=\frac{371.5-588}{\sqrt{9506}}=-2.22
\end{aligned}
$$

Choosing a two-sided alternative hypothesis, the corresponding p-value is $2 P(Z<-2.22)=0.0264$. Thus, we reject the null hypothesis at the $5 \%$ significance level.

## Alternative 3.

$$
\begin{aligned}
E(T) & =\frac{n(n+1)}{4}=\frac{44(44+1)}{4}=495 \\
\operatorname{Var}(T) & =\frac{n(n+1)(2 n+1)}{24}=\frac{44(44+1)(2 \cdot 44+1)}{24}=7342.5, \\
Z_{o b s} & =\frac{T-E(T)}{\sqrt{\operatorname{Var}(T)}}=\frac{371.5-495}{\sqrt{7342.5}}=-1.44
\end{aligned}
$$

[^1]Choosing a two-sided alternative hypothesis, the corresponding p -value is $2 P(Z<-1.44)=0.1498$. Thus, we do not reject the null hypothesis at the $5 \%$ significance level.


[^0]:    ${ }^{1}$ It is a sufficient, but not necessary, condition that both the "Before" and "After" variables are drawn from normally distributed populations. It does not matter whether the population variances of "Before" and "after" are equal or not.
    ${ }^{2}$ The teaching in this course does not include formal tests of normality. It is acceptable to reach another conclusions, but all conclusions should be substantiated.

[^1]:    ${ }^{3}$ Calculations are not necessary, but the $95 \%$ confidence interval for the difference can be calculated as follows:

    $$
    \begin{gathered}
    \bar{x}_{1}-\bar{x}_{2} \pm t_{n_{1}+n_{2}-2,0.25} \sqrt{\frac{s_{p}^{2}}{n_{1}}+\frac{s_{p}^{2}}{n_{2}}} \\
    6.956545-6.198018 \pm 2.021 \sqrt{\frac{0.880853}{21}+\frac{0.880853}{21}} \\
    0.758527 \pm 0.58536036 \\
    (0.173,1.344) .
    \end{gathered}
    $$

    Except for round-off errors, this corresponds to the $95 \%$ confidence interval reported in the "diff" line of the Stata output.

