

**EXAM IN HMET4100, 7 DECEMBER 2021:
SUGGESTED SOLUTIONS AND COMMENTS**

VERSION INTENDED FOR CANDIDATES

Knut R. Wangen

k.r.wangen@medisin.uio.no

December 30, 2021

INTRODUCTION

The exam was arranged digitally on the Inspira platform and consisted of 9 questions. Four 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 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: [75-90], B: [65-74], C: [50-64], D: [41-49], E: [30-40], F: [0-29].

Explanation of grades. After the grades are published, the candidates will have access to their individual scores by logging on to Inspira. The present suggested solution, which will be distributed to the candidates, explains the general principles for how scores and grades were awarded.

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

Informal feedback. A candidate who wishes to receive additional feedback may ask the course coordinator for an informal talk (phone call or Zoom meeting). The candidate must then email the course coordinator 2-3 suggested times with at least two workdays notice. The suggested times should be within normal office hours (0900h-1700h) in the period 5–14 January 2022.

SUGGESTED SOLUTIONS

QUESTION 1

$$\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}} \Rightarrow 0.334 \pm Z_{\alpha/2} \sqrt{\frac{0.334(1-0.334)}{493}} \Rightarrow \underline{\underline{(0.299, 0.369)}}$$

QUESTION 2

$$n = \frac{Z_{\alpha/2}^2 \sigma^2}{ME^2} = \frac{1.645^2 \cdot 55^2}{1.25^2} \approx \underline{\underline{96}}$$

The answer is rounded up to the nearest integer.

QUESTION 3

We have a binomial distribution with $n = 8$ and $p = 1/3$, that is, $p \approx 0.333$.

Solution using Appendix Table 2: From the complement rule, we have $P(X \geq 6) = 1 - P(X \leq 5)$. For $p = 0.30$, $P(X \leq 5) = 0.989$, so that $P(X \geq 6) = 0.011$. For $p = 0.35$, $P(X \leq 5) = 0.975$, so that $P(X \geq 6) = 0.025$. This means that the answer must be between 0.011 and 0.025, and the only alternative satisfying this requirement is 0.020.

Solution obtained by calculating individual probabilities, when $p = 1/3$: $P(X \geq 6) = P(X = 6) + P(X = 7) + P(X = 8) = 0.0171 + 0.0024 + 0.0002 = \underline{\underline{0.020}}$.

QUESTION 4

From Appendix Tables 7a and 7b, we have $P(\chi^2 > 37.652) = 0.05$ and $P(\chi^2 < 14.611) = 0.05$. This implies that $P(14.611 < \chi^2 < 37.652) = \underline{\underline{0.10}}$.

QUESTION 5

$$P(X \geq 3) = P(X \leq 2) = 1 - 0.0884 = \underline{\underline{0.9116}}$$

QUESTION 6

a) The upper test assumes that the unknown population variances are equal. For a two-sided test, the p-values is 0.0045. Hence, we reject the null hypothesis (difference = zero) at the 5% significance level.

The lower test does not assume that the unknown population variances are equal. For a two-sided test, the p-values is 0.0679. Hence, we *do not* reject the null hypothesis at the 5% significance level.

b) The tests differ regarding the assumption on the unknown population variances. This implies that the formulas for the standard errors are different.

In the upper test (variances assumed equal), the t-statistic has $n_1 + n_2 - 2 = 12 + 6 - 2 = 16$ degrees of freedom. In the lower test (variances not necessarily equal), the t-statistic's degrees of freedom are given by the formula (15) in the "collection of selected formulas". In this case, the degrees of freedom are 5.1199.

c) From the table we see that the sample standard deviation is 21.58844 in the combined intervention groups (Interv) and 139.6324 in the control group (Group3). The respective variances are 466 and 19497, that is, the control group has a variance nearly 42 times higher than the combined intervention

groups. It does not seem appropriate to assume that the variances are equal, and, hence, the lower test seems most appropriate.

QUESTION 7

We first calculate $U = n_1n_2 + \frac{n_1(n_1+1)}{2} - R_1 = 12 \cdot 6 + \frac{12(12+1)}{2} - 137 = 13$, $E(U) = \frac{n_1n_2}{2} = \frac{12 \cdot 6}{2} = 36$, and $Var(U) = \frac{12 \cdot 6(12+6+1)}{12} = 114$.

We then calculate test statistic, $Z_{obs} = \frac{U-E(U)}{\sqrt{Var(U)}} = \frac{13-36}{\sqrt{114}} = -2.15$.

The p-value is $P(Z < -2.15) + P(Z > 2.15) = 0.0158 + 0.0158 = 0.0316$. We reject the null hypothesis on the 5% significance level. Comment: The normal approximation usually requires that there are at least 10 observations in each sample, and in this case we have only 6 observations in the control group.

QUESTION 8

The null hypothesis is that the median reduction in headache activity are equal in all three groups. The alternative hypothesis is that at least one median is different from the others.

We find the observed test statistics from the formula:

$$W_{obs} = \frac{12}{n(n+1)} \sum_{i=1}^K \frac{R_i^2}{n_i} - 3(n+1) = 5.696.$$

With $K = 3$ groups, the critical value is $W_{critical} = \chi_{2,0.05}^2 = 5.991$.

We do not reject the null hypothesis at the 5% significance level.

QUESTION 9

a)

The F-ratio in the table, 5.19, should be compared to the critical value at the 5% significance level, which is $F_{K,n-K,\alpha} = F_{2,15,0.05} = 3.682$. The F-ratio is higher than the critical value, so we reject the null hypothesis.

b) Several answers are possible. Some of these aspects should be considered:

Small sample sizes generally weigh in favor of non-parametric tests, unless other existing information indicate that the populations have normal distributions. The box plots may indicate that all groups have slightly left skewed distributions, but as there are no extreme outliers this skewness does not need to pose a problem when using tests based on normally distributed populations.

We kept the null hypothesis for the Kruskal-Wallis test, while we rejected the null hypothesis for the ANOVA test. Note that the ANOVA test relies on the assumption of normal populations and an assumption that the population variances are identical. The Kruskal-Wallis test would generally be considered more appropriate when we have small samples and unknown population distributions.

Merging the intervention groups (questions 6 and 7) imposes a restriction that may not be valid and which is difficult to test in small samples (i.e.,

that these two groups are drawn from the same distribution). The box plots suggest that this restriction may be valid — or at least that the difference between the two intervention groups are negligible when compared to the control group.

The t-test assuming equal variances is clearly not appropriate, given the differences in variances mentioned in the answer to question 6, but also because the box plots reveal that group 3 (the control group) has a much larger variability than groups 1 and 2 (intervention groups).

For the t-test with unequal variances we kept the null hypothesis, while for the Mann-Whitney test we rejected the null hypothesis. The most important difference between them is that the t-test assumes that the populations have normal distributions while the Mann-Whitney test does not. The Mann-Whitney test in question 7 is approximately valid when there are at least 10 observations in each group, and in this case one group had fewer observations. We do not have a box plot for the two intervention groups combined, and this makes it more difficult to assess normality.

It can be useful to ask which test one should choose *a priori*, that is, before the samples became available. If we do not know whether the populations are normally distributed, or whether the two interventions groups are identical so that they can be merged, it seems natural to choose the Kruskal-Wallis test.