

**EXAM IN HMET4100, 12 DECEMBER 2019:
SUGGESTED SOLUTIONS AND COMMENTS**

VERSION INTENDED FOR CANDIDATES

Knut R. Wangen

k.r.wangen@medisin.uio.no

December 31, 2019

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 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	14	29.6%
B	13	25.0%
C	16	30.8%
D	3	5.8%
E	4	7.7%
F	2	3.8%

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. 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 web-page (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 the period 8–31 January 2020.

SUGGESTED SOLUTIONS

QUESTION 1

Let X denote the number of successes and $p = 0.75$ denote the probability of success in each trial. We can then let Y denote the number of failures and $q = 1 - p = 0.25$ denote the probability of failure in each trial.

$$\begin{aligned} P(X \leq 7) &= P(X = 0) + \dots + P(X = 7) \\ &= P(Y = 14) + \dots + P(Y = 7) = 1 - P(Y \leq 6) \\ &= 1 - 0.962 \approx \underline{\underline{0.038}}. \end{aligned}$$

QUESTION 2

$$P(Y > -1.492) = 1 - P(Y < -1.492) = 1 - P(Y > 1.492) = 1 - 0.01 = \underline{\underline{0.99}}.$$

The second equality follows from the fact that the t -distribution is symmetric about zero.

QUESTION 3

Using the formula

$$\hat{p} \pm Z_{\alpha/2} \sqrt{\frac{\hat{p}(1 - \hat{p})}{n}},$$

and inserting values

$$0.203 \pm 1.645 \sqrt{\frac{0.203(1 - 0.203)}{763}},$$

we obtain the confidence interval $\underline{\underline{(0.179, 0.227)}}$.

QUESTION 4

$$n = \frac{Z_{\alpha/2}^2 \sigma^2}{ME^2} = \frac{1.645^2 \cdot 50}{1^2} = 135.3,$$

that is rounded up to $\underline{\underline{136}}$.

QUESTION 5

(Several solutions are possible.) The hypotheses state relationships between two expectations but the text mentions “a sample”. One possible interpretation is that the same units of observation (e.g., persons) are observed twice (dependent samples). Another possible interpretation is that there are two independent samples that have in total 1300 observations and are about the same sample size. In both interpretations, the central limit theorem can be used to justify that the test statistic is approximately normally distributed.

a) In a two-sided test, the p-value is $P(Z > |-2.39|) = P(Z < -2.39) + P(Z > 2.39) = 2 \cdot 0.0084 = 0.0168$. The test conclusion is that we do not reject the null hypothesis at the 1% significance level, because the p-value is greater than the significance level.

b) There are two possible one-sided alternatives, “ $\mu_x > \mu_y$ ” and “ $\mu_x < \mu_y$ ”. One of them will have the p-value $P(Z < -2.39) = 0.0084$ and lead us to reject the null hypothesis because the p-value is less than the significance level. The other one-sided alternative hypothesis will have the p-value $P(Z > -2.39) = 1 - 0.0084 = 0.9916$ and lead us to keep the null hypothesis.

QUESTION 6

The null hypothesis was of the form “ $\mu_{\text{control}} = \mu_{\text{treatment}}$ ” or “ $\mu_{\text{control}} - \mu_{\text{treatment}} = 0$ ”. That is, the difference in mean sleep quality is assumed to be zero.

The test results are from a two-sample t-test where the population variances are unknown and not assumed equal. If we let x and y denote the control and treatment group, respectively, the test statistic is

$$T = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_x} + \frac{s_y^2}{n_y}}} = \frac{4.571038 - 4.781609}{\sqrt{\frac{1.069698^2}{366} + \frac{1.206152^2}{783}}} = -2.983.$$

The number of degrees of freedom is 796.451 which means that the critical values at the 1% significance level are 2.576 and -2.576 (when using infinity as the degrees of freedom, that is, using the standard normal distribution)¹. We reject the null hypothesis and accept the alternative hypothesis because $|Z_{\text{obs}}| = 2.983 > 2.576$.

QUESTION 7

(several answers are possible.) The test in question 6 relies on the assumption that the samples are drawn from two normally distributed populations. Further, the two population variances are considered to be unknown and generally not equal. Given these assumptions and the large samples sizes, with 366 and 783 observations, the student t -distribution is approximately equal to the standard normal distribution.

¹Also the critical values from the t -distribution with 100 degrees of freedom are acceptable, that is, ± 2.626 .

However, the variable at hand is discrete and cannot be normally distributed in a strict sense. The histograms indicate that the sample distributions are skewed to the right, and not symmetric as the normal distribution. The skewness can also be seen in the box plots, where the medians are further from the maximum values than from the minimum values. For the control group, the median (50% percentile) is equal to the first quartile (25% percentile). This is possible because the variable is discrete and not continuous, and it indicates that the control's sample distribution has a relatively high density at the left side of the median. Summing up, the sample distributions appear to be non-normal and, given the large sample sizes, this suggests that the populations are non-normal.

If the samples sizes are sufficiently large, the central limit theorem ensures that the sample means and, hence, the test statistic are approximately normal even if the population distributions should be non-normal. The sample sizes in this case will often be considered large enough for us to expect that the test statistic is approximately normal. Moreover, the sleep quality variable has a limited range (1–10), and with the sample medians close to the middle of the range it is unlikely that outliers should notably reduce the accuracy of the normal approximation. Thus, it seems reasonable to consider the test result to be reliable.²

QUESTION 8

There are three groups, each with a small number of observations. The box plots indicate that the distributions are non-normal. Thus, it is appropriate to use the non-parametric Kruskal-Wallis test.

The null hypothesis is “ H_0 : The medians in the three groups are equal” while the alternative hypothesis is “ H_1 : Not all the three medians are equal”.

If we choose a 5% significance level, the critical value is $W_{critical} = \chi^2_{2,0.05} = 5.991$. We then calculate the test statistic

$$\begin{aligned} W &= \frac{12}{n(n+1)} \sum_{i=1}^K \frac{R_i^2}{n_i} - 3(n+1) \\ &= \frac{12}{40(40+1)} \left(\frac{288^2}{12} + \frac{315^2}{13} + \frac{217^2}{15} \right) - 3(40+1) = 6.395. \end{aligned}$$

The test conclusion is that we reject the null hypothesis and accept the alternative hypothesis, that is, the three medians are not equal. This result suggests that the three treatments do not have the same effect.³

²If non-parametric methods are considered, the appropriate alternatives are the Mann-Whitney U Test and the Wilcoxon Rank Sum Test (the Sign Test and the Wilcoxon Signed Rank Test are not appropriate for independent samples).

³If we choose the 1% significance level the critical value is 9.210 and we would not reject the null hypothesis.

QUESTION 9

In order to calculate the test statistic when using the normal approximation, we first find

$$E(T) = \mu_T = \frac{n(n+1)}{4} = \frac{20 \cdot 21}{4} = 105,$$

$$Var(T) = \sigma_T^2 = \frac{n(n+1)(2n+1)}{24} = \frac{20(20+1)(2 \cdot 20+1)}{24} = 717.5,$$

and $T = \min(T+, T-) = \min(160.5, 49.5) = 49.5$. We then calculate the observed test statistic

$$Z_{obs} = \frac{T - \mu_T}{\sigma_T} = \frac{49.5 - 105}{\sqrt{717.5}} = -2.072$$

The critical value at the 1% significance level is 2.576. We do not reject the null hypothesis because $|Z_{obs}| = 2.072 < 2.576$.⁴

⁴The text states explicitly that the normal approximation should be used. The problem can be solved by comparing $T = 49.5$ to a critical value obtained from Table 10 in the textbook. For $n = 20$, the critical value is 38 (note that Table 10 presents critical values for one-sided alternative hypotheses). We keep H_0 because the observed test statistic (49.9) is higher than the critical value (38).