Technical University of Denmark *Page 1 of 21 pages.*

Written examination: 15. December 2019

Course name and number: **Introduction to Statistics (02402)**

Duration: 4 hours

Aids and facilities allowed: All

The questions were answered by

_____ _____ _____

(student number) (signature) (table number)

This exam consists of 30 questions of the "multiple choice" type, which are divided between 12 exercises. To answer the questions, you need to fill in the "multiple choice" form (6 separate pages) on CampusNet with the numbers of the answers that you believe to be correct.

5 points are given for a correct "multiple choice" answer, and -1 point is given for a wrong answer. ONLY the following 5 answer options are valid: 1, 2, 3, 4, or 5. If a question is left blank or an invalid answer is entered, 0 points are given for the question. Furthermore, if more than one answer option is selected for a single question, which is in fact technically possible in the online system, 0 points are given for the question. The number of points needed to obtain a specific mark or to pass the exam is ultimately determined during censoring.

The final answers should be given by filling in and submitting the form online via CampusNet. The table provided here is ONLY an emergency alternative. Remember to provide your student number if you do hand in on paper.

Exercise	I.1	I.2	I.3	II.1	III.1	IV.1	V.1	V.2	V.3	V.4
Question	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Answer										

Exercise	VI.1	VI.2	VII.1	VII.2	VII.3	VII.4	VII.5	VII.6	VIII.1	VIII.2
Question	(11)	(12)	(13)	(14)	(15)	(16)	(17)	(18)	(19)	(20)
Answer										

Exercise	VIII.3	IX.1	IX.2	X.1	X.2	XI.1	XI.2	XII.1	XII.2	XII.3
Question	(21)	(22)	(23)	(24)	(25)	(26)	(27)	(28)	(29)	(30)
Answer										

The exam paper contains 21 pages. Continue on page 2

Multiple choice questions: *Note that in each question, one and only one of the answer options is correct. Furthermore, not all the suggested answers are necessarily meaningful. Always remember to round your own result to the number of decimals given in the answer options before you choose your answer.*

Exercise I

A biodynamic farm wants to degrade its biomass residues into bio liquid to be used for renewable energy production. In an experiment, the farmers used 10 liter reaction containers to assess the efficiency of the biomass conversion. Varying amounts of an enzymatic cocktail were added to each of the containers, and the mixtures were left for three days of reaction time. Afterwards, the volumes of produced bio liquid were determined.

A simple linear regression model of the form $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$ was established, in order to investigate the relationship between the amount of enzyme added (`enzyme`, in ml) and the bio liquid yield (`liquid`, in dl). The R output from fitting the model can be seen below:

```
##
## Call:
## lm(formula = liquid ~ enzyme)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -9.8103 -4.3885 -0.0775  4.3672  9.2489
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)    9.325      2.291   4.070 0.000653 ***
## enzyme         1.956      0.196   9.982 5.42e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 5.438 on 19 degrees of freedom
## Multiple R-squared:  0.8398,Adjusted R-squared:  0.8314
## F-statistic: 99.64 on 1 and 19 DF,  p-value: 5.419e-09
```

Question I.1 (1)

Given the R output above, what is the sample size n?

1 ☐ 20

2 ☐ 19

3 ☐ 21

4 ☐ 1

5 ☐ The sample size cannot be determined from this R output.

Question I.2 (2)

In the experiment, the average amount of enzymatic cocktail used in a reaction container was $\bar{x} = 10$ ml. Compute the average bio liquid yield, $\bar{y}$.

1 ☐ $\bar{y} = 28.9$ dl

2 ☐ $\bar{y} = 9.3$ dl

3 ☐ $\bar{y} = 2.0$ dl

4 ☐ $\bar{y} = 19.6$ dl

5 ☐ $\bar{y} = 24.1$ dl

Question I.3 (3)

Which of the statements below does not represent a necessary assumption for a simple linear regression model?

1 ☐ The errors ε_i are independent.

2 ☐ The errors ε_i are identically distributed.

3 ☐ The outcomes Y_i are identically distributed.

4 ☐ The outcomes Y_i are independent.

5 ☐ The outcomes Y_i and the errors ε_i have the same variance.

Continue on page 4

Exercise II

In connection with the examination in an introductory statistics course, one wants to examine whether students, who have been enrolled in the study program for one year, perform differently than students who have been enrolled for two years. The exam score is calculated as a number between -30 and 150 by the rules:

- there are 30 questions in total,
- -1 point is given for a wrong answer,
- 5 points are given for a correct answer,
- only one answer can be given to each question.

Two samples consisting of exam scores have been collected randomly from the students: One from students who are in their first year (`x`), and one from students who are in their second year (`y`).

The samples each contains 50 observations, and their means are $\bar{x} = 84.0$ and $\bar{y} = 86.6$, respectively. The following simulations and calculations are carried out in R:

```
k <- 10000

simxsamples <- replicate(k, sample(x, replace = TRUE))
simysamples <- replicate(k, sample(y, replace = TRUE))
simmeandifs <- apply(simxsamples, 2, mean) - apply(simysamples, 2, mean)

quantile(simmeandifs, c(0.05, 0.95))

##     5%    95%
## -15.12   9.87

quantile(simmeandifs, c(0.025, 0.975))

##   2.5%  97.5%
## -17.26  12.42

quantile(simmeandifs, c(0.005, 0.995))

##   0.5%  99.5%
## -22.04  17.32
```

Question II.1 (4)

The null hypothesis

$$H_0 : \mu_X = \mu_Y$$

is to be tested at significance level $\alpha = 5\%$, without making assumptions about the distribution of the scores in the two samples. Which of the following answers is correct? (Both the conclusions and argument must hold).

1 ☐ The null hypothesis is not rejected, as $0 \in [-17.26, 12.42]$. Hence, a significant difference cannot be detected.

2 ☐ The null hypothesis is not rejected, as $2.6 \in [-15.12, 9.87]$. Hence, a significant difference cannot be detected.

3 ☐ The null hypothesis is rejected, as $0 \in [-17.26, 12.42]$. Hence, it can be established that students in their first year perform better that students in their second year.

4 ☐ The null hypothesis is rejected, as $0 \notin [-22.04, 17.32]$. Hence, it can be established that students in their first year perform better that students in their second year.

5 ☐ The null hypothesis is not rejected, as $2.6 \in [-22.04, 17.32]$. Hence, it can be established that students in their second year perform better that students in their first year.

Continue on page 6

Exercise III

In a hospital, a group of patients are randomly selected. They receive a questionnaire about the hospital's service, both when they are admitted and when they leave the hospital. In both questionnaires, the patients are asked to indicate their satisfaction with the hospital's service on a continuous scale from 0 to 1. Subsequent analysis of data reveals that both series of measurements of service satisfaction can be assumed to be normally distributed. Which of the following 5 tests is most suitable for a comparison of the service assessment upon hospitalization and when leaving the hospital?

Question III.1 (5)

1 ☐ A χ^2-test in a contingency table

2 ☐ A one-way analysis of variance

3 ☐ A t-test with two independent samples

4 ☐ A paired t-test

5 ☐ A regression analysis

Continue on page 7

Exercise IV

Question IV.1 (6)

Assume that the random variable $X \in [0, 1]$ follows a distribution with density function $f(x) = 2x$ for $x \in [0, 1]$, and thus has the distribution function $F(x) = x^2$. Which of the following pieces of R code simulates outcomes of the random variable X?

1 ☐ `2 * runif(k)`

2 ☐ `rchisq(k, df = 1)`

3 ☐ `runif(k)^2`

4 ☐ `rchisq(k, df = k - 1)`

5 ☐ `sqrt(runif(k))`

Continue on page 8

Exercise V

A plastic manufacturer wants to determine if there is a difference in the quality of plastic produced with materials from different suppliers (`Supplier`). In the production, a particular measured variable Y (`y`) is known to determine the quality of the produced plastic. Higher values of Y indicate higher quality of the produced plastic. The table below shows values of Y collected from separate production runs with materials from 5 different suppliers. Subsequently, output from the analysis that was run in R by the company's engineers is shown.

Supplier A	Supplier B	Supplier C	Supplier D	Supplier E
9.9	8.7	8.3	10.4	7.7
10.5	10.3	10.7	12.1	11.7
8.2	6.1	8.7	11.5	10.1
7.7	7.6	9.5	11.2	9.0

```
anova(lm(y ~ Supplier))

## Analysis of Variance Table
##
## Response: y
##           Df Sum Sq Mean Sq F value Pr(>F)
## Supplier   4   20.9    5.23    2.77  0.066 .
## Residuals 15   28.3    1.89
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Question V.1 (7)

Given the model used in the analysis, what is the estimate of the expected value $\mathrm{E}(Y_{\mathrm{D},i})$ for supplier D?

1 ☐ 10.3

2 ☐ 10.5

3 ☐ 10.8

4 ☐ 11.3

5 ☐ 11.5

Question V.2 (8)

Which of the following answers most accurately describes the hypothesis tested in the R output above?

1 ☐ The hypothesis $\alpha_i = 1$ for all $i = 1, 2, 3, 4, 5$, in a model of the form $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$.

2 ☐ The hypothesis $\alpha_i = 0$ for all $i = 1, 2, 3, 4, 5$, in a model of the form $Y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$.

3 ☐ The hypothesis $\beta_0 = 1$ in a model of the form $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$.

4 ☐ The hypothesis $\beta_1 = 0$ in a model of the form $Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i$.

5 ☐ None of the above answers describe the hypothesis that is tested.

Question V.3 (9)

Use the significance level $\alpha = 5\%$. Is there a significant difference in the quality of the plastic produced with the materials from the 5 suppliers? (Both the conclusion and argument must hold).

1 ☐ A significant difference in quality cannot be detected, as the p-value is above the significance level.

2 ☐ A significant difference in quality can be detected, as the p-value is under the significance level.

3 ☐ A significant difference in quality cannot be detected, as the p-value is under the significance level.

4 ☐ A significant difference in quality can be detected, as the p-value is above the significance level.

5 ☐ None of the above conclusions are correct.

Question V.4 (10)

How much of the total variation cannot be explained by the model?

1 ☐ $\frac{16.7}{28.3+20.9} = 33.9\%$

2 ☐ $\frac{20.9}{28.3+20.9} = 42.5\%$

3 ☐ 6.6%

4 ☐ $\frac{28.3}{28.3+20.9} = 57.5\%$

5 ☐ $\frac{1.89}{1.89+5.23} = 26.5\%$

Continue on page 10

Exercise VI

The value of X has been measured for 5 individuals in Group 1 and 10 individuals in Group 2, respectively. It can be assumed that the observations in both groups are normally distributed and that all the observations are mutually independent. The variances in the two groups are allowed to be different. One would like to test the hypothesis that the two groups have the same mean (against the alternative that the means are different). The test is performed at a 5% significance level.

Question VI.1 (11)

By a comparison with the usual test statistic, which of the following quantiles can easily be used in order to determine whether there is a significant difference between the two means?

1 ☐ The 0.025 quantile of the relevant t-distribution.

2 ☐ The 0.05 quantile of the relevant t-distribution.

3 ☐ The 0.95 quantile of the standard normal distribution.

4 ☐ The 0.90 quantile of the standard normal distribution.

5 ☐ The 0.50 quantile of the standard normal distribution.

Question VI.2 (12)

The sample mean and standard deviation in Group 1 are, $\bar{x}_1 = 1.99$ and $s_1 = 0.58$, while the corresponding numbers for Group 2 are, $\bar{x}_2 = 1.14$ and $s_2 = 0.84$. The variances in the two groups are assumed to be different. In this case, the test statistic for the above test is:

1 ☐ $t_{\text{obs}} = 4.3$

2 ☐ $t_{\text{obs}} = 1.9$

3 ☐ $t_{\text{obs}} = 2.3$

4 ☐ $t_{\text{obs}} = 6.2$

5 ☐ None of the above possibilities.

Continue on page 11

Exercise VII

A research project involves collecting insects by driving predefined trips with a net on the roof of a car. After the trip, the collected insects are sent to the university for counting.

Question VII.1 (13)

Which of the following distributions is presumably best for describing the number of insects in a net?

1 ☐ An exponential distribution

2 ☐ A binomial distribution

3 ☐ A normal distribution

4 ☐ A hypergeometric distribution

5 ☐ A Poisson distribution

Question VII.2 (14)

A total of four trips are planned on the same stretch of road, and it is assumed that the variance of the number of insects, σ^2, is the same on each of the four trips. Furthermore, the results of the four trips are assumed to be independent. What is the variance of the total number of insects captured on the four trips?

1 ☐ $16\sigma^2$

2 ☐ $4\sigma^2$

3 ☐ $\sigma^2/4$

4 ☐ 4σ

5 ☐ $\sigma^2/2$

The four planned trips are carried out, and the captured insects divided into two types: small and large insects. The result of the counting is shown in the contingency table below.

	Trip 1	Trip 2	Trip 3	Trip 4	Total
Small insects	178	242	126	87	633
Large insects	26	59	30	8	123
Total	204	301	156	95	756

Question VII.3 (15)

Looking at the overall result (all four trips combined), which of the following is a 95% confidence interval for the proportion of large insects?

1 ☐ $[0.14; 0.19]$

2 ☐ $[0.81; 0.87]$

3 ☐ $[0.16; 0.23]$

4 ☐ $[0.09; 0.23]$

5 ☐ $[0.81; 0.86]$

Question VII.4 (16)

There are special reasons for examining whether the proportion of large insects can be assumed to be the same on Trip 1 and Trip 2. What is the p-value and the conclusion at significance level $\alpha = 5\%$, for a test investigating whether the proportion of large insects differs between Trip 1 and Trip 2?

1 ☐ The p-value is 0.043, and a difference can therefore be established.

2 ☐ The p-value is 0.03 and a difference can therefore be established.

3 ☐ The p-value is 0.060 and a difference can therefore be established.

4 ☐ The p-value is 0.043 and therefore no difference can be established.

5 ☐ The p-value is 0.060 and therefore no difference can be established.

In the following questions, we look at data from all four trips, in order to test whether the distribution between large and small insects can be assumed to be the same on all trips.

Question VII.5 (17)

In order to perform the statistical test, the expected number of insects in each cell, under the null hypothesis, must be calculated. What is the expected number of large insects on Trip 3?

1 ☐ 130.6

2 ☐ 4.9

3 ☐ 105.5

4 ☐ 25.4

5 ☐ 32.1

Question VII.6 (18)

The usual test statistic for examining the difference between the distribution of the number of insects on the four trips is calculated to be 9.6127. What is the p-value and the corresponding conclusion at significance level $\alpha = 5\%$?

1 ☐ The p-value is 0.022, so no difference can be detected.

2 ☐ The p-value is 0.087, hence there is a difference.

3 ☐ The p-value is 0.087, so no difference can be detected.

4 ☐ The p-value is 0.022, hence there is a difference.

5 ☐ The p-value is 0.045, hence there is a difference.

Continue on page 14

Exercise VIII

The amount of detergent necessary for washing laundry typically depends on several factors. In this context, the relationship between washing efficiency (`efficiency`), water hardness (`hardness`), and the amount of detergent used (`detergent`) is to be investigated using the following multiple linear regression model:

$$efficiency_i = \beta_0 + \beta_1 \cdot hardness_i + \beta_2 \cdot detergent_i + \varepsilon_i \,,$$

where the ε_i are independent and $N(0, \sigma^2)$-distributed. R output from the model is shown below:

```
##
## Call:
## lm(formula = efficiency ~ hardness + detergent)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -3.9022 -1.4491 -0.5854  1.4225  5.3286
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  -1.5892     3.6590  -0.434   0.6695
## hardness     -2.1981     0.8958  -2.454   0.0252 *
## detergent     3.0239     0.4961   6.095 1.19e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.496 on 17 degrees of freedom
## Multiple R-squared:  0.7322,Adjusted R-squared:  0.7006
## F-statistic: 23.23 on 2 and 17 DF,  p-value: 1.371e-05
```

Question VIII.1 (19)

Look at the R output above. Which of the following statements is correct, given a significance level of $\alpha = 1\%$?

1 ☐ Water hardness appears to have a significant effect on washing efficiency, while the amount of detergent used does not.

2 ☐ The effect of water hardness on washing efficiency is not significant, because the p-value is greater than 0.01.

3 ☐ Both water hardness and the amount of detergent used are significant, because the p-values are less than 0.05.

4 ☐ Neither water hardness nor the amount of detergent used appear to be significant, because the p-values are less than 0.05.

5 ☐ The model intercept is significant, because the p-value of 0.6695 is greater than 0.01.

Question VIII.2 (20)

Look at the same R output above. What effect does an increase of two units of detergent have on expected washing efficiency? Assume water hardness to be constant.

1 ☐ The expected washing efficiency increases by 3.02 units.

2 ☐ The expected washing efficiency decreases by 2.20 units.

3 ☐ The expected washing efficiency decreases by 4.40 units.

4 ☐ The expected washing efficiency increases by 6.05 units.

5 ☐ The expected washing efficiency remains constant.

Question VIII.3 (21)

Give an estimate of the variance σ^2 based on the R output above.

1 ☐ $\hat{\sigma}^2 = 23.23$

2 ☐ $\hat{\sigma}^2 = 0.7006$

3 ☐ $\hat{\sigma}^2 = 2.496$

4 ☐ $\hat{\sigma}^2 = 0.7322$

5 ☐ $\hat{\sigma}^2 = 6.230$

Continue on page 16

Exercise IX

The temperature in a refrigerator was measured at 12 o'clock on randomly selected days during the month of July. The following observations were measured (in degrees celsius) and loaded into R in the vector `x`:

```
x <- c(6.5, 5.7, 1.2, 0.2, 7.0, 3.3)
```

The observations are assumed to be normally distributed and mutually independent.

Question IX.1 (22)

Compute the usual test statistic for testing the hypothesis that the mean is 3.0 degrees.

1 ☐ $t_{\text{obs}} = 0.84$

2 ☐ $t_{\text{obs}} = 0.20$

3 ☐ $t_{\text{obs}} = 2.41$

4 ☐ $t_{\text{obs}} = 3.01$

5 ☐ $t_{\text{obs}} = 1.99$

Question IX.2 (23)

Determine a 90% confidence interval for the variance of the refrigerator temperature.

1 ☐ $[2.9, 51.5]$

2 ☐ $[3.2, 49.2]$

3 ☐ $[3.7, 35.7]$

4 ☐ $[3.9, 33.7]$

5 ☐ $[4.1, 31.7]$

Continue on page 17

Exercise X

The fuel consumption of four different tractors was investigated in connection with three different tasks, and the following results were obtained (in litres per hectare):

	Tractor A	Tractor B	Tractor C	Tractor D
Task 1	8.1	8.3	9.2	8.7
Task 2	12.2	11.8	14.2	13.1
Task 3	8.9	9.1	7.3	8.2

Question X.1 (24)

Which of the following methods is best suited to investigate whether there is a difference in the fuel consumption of the different tractors?

1 ☐ A test in a multiple linear regression model

2 ☐ A χ^2-test in a contingency table

3 ☐ A two-sample t-test

4 ☐ One-way analysis of variance

5 ☐ Two-way analysis of variance

Question X.2 (25)

The median fuel consumption (expressed in litres per hectare) for Task 3 is:

1 ☐ 7.3

2 ☐ 8.2

3 ☐ 8.375

4 ☐ 8.55

5 ☐ 9.1

Continue on page 18

Exercise XI

Let $X \sim N(0, \sigma^2)$ and define the random variable Y by $Y = e^X$.

Question XI.1 (26)

What is $P(Y > 1)$?

1 ☐ 0.84

2 ☐ 0.5

3 ☐ 0.025

4 ☐ 0.16

5 ☐ 0.95

Question XI.2 (27)

What is the variance of Y?

1 ☐ $e^{\sigma^2/2}$

2 ☐ $e^{2+\sigma^2}(e^{1/2} - 1)$

3 ☐ $e^{\sigma^2}(e^{\sigma^2} - 1)$

4 ☐ $e^{2+\sigma^2}(e^{\sigma^2} - 1)$

5 ☐ e^{σ^2}

Continue on page 19

Exercise XII

A manufacturer would like to examine the quality of its production facilities. A random sample, consisting of observed times between the production of faulty elements, was collected from the production plant. The values are in hours and loaded into R with the following code:

```
x <- c(39.5, 59.7, 42.1, 13, 3.6, 10.9, 61.6, 1, 17.8, 5,
       24.3, 21, 4.2, 21.1, 78.9, 11.1, 6.6, 0.3, 9.2, 10.4)
```

Question XII.1 (28)

Use the book's definition of sample quantiles to determine the *IQR* (*"Inter Quartile Range"*) of the sample.

1 ☐ $IQR = 69.6$

2 ☐ $IQR = 26.1$

3 ☐ $IQR = 58.35$

4 ☐ $IQR = 6.25$

5 ☐ $IQR = 16.05$

Question XII.2 (29)

It has been decided that the plant must be stopped and repaired if the time between the faults becomes too short. To avoid making assumptions regarding the distribution of the time between faults, one would like to construct a non-parametric 95% bootstrap confidence interval for the median. Which of the following R codes determines this interval correctly?

1 ☐
```
simsamples <- replicate(10000, sample(x, replace = TRUE))
quantile(apply(simsamples, 2, mean), c(0.05, 0.95))
```

2 ☐
```
simsamples <- replicate(10000, sample(x, replace = FALSE))
quantile(apply(simsamples, 2, mean), c(0.025, 0.975))
```

3 ☐
```
simsamples <- replicate(10000, sample(x, replace = FALSE))
quantile(apply(simsamples, 2, median), c(0.05, 0.95))
```

4 ☐
```
simsamples <- replicate(10000, sample(x, replace = TRUE))
quantile(apply(simsamples, 2, median), c(0.025, 0.975))
```

5 ☐
```
simsamples <- replicate(10000, sample(x, replace = TRUE))
quantile(apply(simsamples, 2, median), c(0.005, 0.995))
```

Question XII.3 (30)

After a repair of the plant, a new sample is collected and entered into R with the code below:

```
y <- c(15.3, 28.2, 53.3, 42, 28.5, 45.3, 40.3, 32.3, 81.1, 29.3,
       82.9, 38.7, 131.5, 24.7, 5.7, 104.3, 30, 31.8, 46.9, 34.9)
```

Subsequently, the following simulations and calculations are carried out:

```
simXsamples <- replicate(10000, rexp(length(x), 1/mean(x)))
simYsamples <- replicate(10000, rexp(length(y), 1/mean(y)))
simDiff <- apply(simXsamples, 2, median) - apply(simYsamples, 2, median)

quantile(simDiff, c(0.005,0.995))

##       0.5%      99.5%
## -50.595024   7.893082

quantile(simDiff, c(0.025,0.975))

##       2.5%      97.5%
## -42.009475   2.646692

quantile(simDiff, c(0.05,0.95))

##           5%          95%
## -37.50105759  -0.01677407
```

Which of the following conclusions is correct based on the R output in this question?

1 ☐ At $\alpha = 1\%$ significance level it may be concluded that there is a significant difference in medians, when no assumptions are made about the distributions of the times.

2 ☐ At $\alpha = 5\%$ significance level it may be concluded that there is no significant difference in medians, under the assumption that the times in both samples are exponentially distributed.

3 ☐ At $\alpha = 10\%$ significance level it may be concluded that there is no significant difference in means, under the assumption that the times in both samples are exponentially distributed.

4 ☐ At $\alpha = 10\%$ significance level it may be concluded that there is a significant difference in means, under the assumption that the times in both samples are normally distributed.

5 ☐ At $\alpha = 1\%$ significance level it may be concluded that there is a significant difference in means, when no assumptions are made about the distributions of the times.

The exam paper is finished. Have a great Christmas vacation!