MAT 2377C
Final exam

17 December 2012 **Professor: Rafal Kulik**
Time: 180 minutes

Student Number: ____________________

Family Name: ____________________

First Name: ____________________

This is a closed book examination.
Only non-programmable and non-graphic calculators are permitted.
Record your answer to each question in the table below.
Your package includes the title page, seven pages with questions, a formula sheet and tables. Note that the binomial table is not given, you will have to use the binomial formula.
Number of questions: **24**.
NOTE: At the end of the examination, hand in only this page. You may keep the questionnaire.

Question	Answer	Question	Answer
1		13	
2		14	
3		15	
4		16	
5		17	
6		18	
7		19	
8		20	
9		21	
10		22	
11		23	
12		24	

GOOD LUCK !!!

Q1. A class consists of 490 female and 510 male students. The students are divided according to their marks

	Passed	Did not pass
Female	430	60
Male	410	100

If one person is selected randomly, the probability that it did not pass given that it is male is (choose the closest answer):

(a) 0.067 (b) 0.123 (c) 0.410

(d) 0.196 (e) None of the preceding

Solution to Q1:

A - 'pass', A^c - did not pass, F - female, F^c - 'not female'=male. To compute: $P(A^c|F^c) = \dfrac{P(A^c \cap F^c)}{P(F^c)} = \dfrac{100/1000}{510/1000} = 0.20$.

Q2. A medical research team wished to evaluate a proposed screening test for Alzheimer's disease. The test was given to a random sample of 450 patients with Alzheimer's disease, in 436 cases the test result was positive. Also, the test was given to a random sample of 500 patients without the disease, only in 5 cases the result was positive. It is known that in the Canada 7.7% of the population aged 65 and over have Alzheimer's disease. Find the probability that a person **has no disease** given that the test was positive (choose the closest answer).

(a) 0.97 (b) 0.88 (c) 0.99

(d) 0.12 (e) None of the preceding

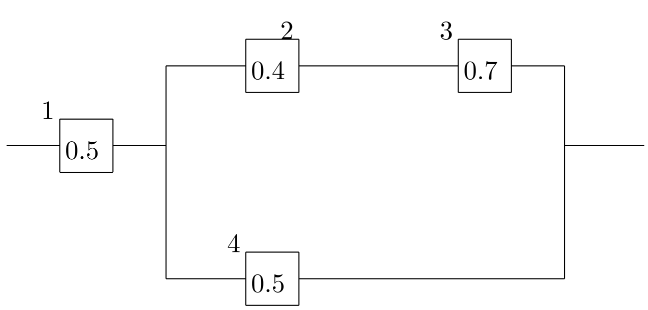
Solution to Q2:

A - test positive, D - a person has disease. Given: $P(A|D) = \dfrac{436}{450}$, $P(A|D^c) = \dfrac{5}{500}$, $P(D) = 0.077$. To find: $P(D^c|A)$ (Bayes' formula):

$$P(D|A) = \frac{P(A|D)P(D)}{P(A|D)P(D) + P(A|D^c)P(D^c)} = 0.88.$$

Hence, $P(D^c|A) = 0.12$

Q3. Consider the following system with four components. We say that it is functional if there exists a path of functional components from left to right. The probability of each component functions is shown. Assume that the components function or fail independently. What is the probability that the system **does not operate**?



(a)	0.32	(b)	0.16	(c)	0.035
(d)	0.68	(e)	None of the preceding		

Solution to Q3:

Call 'Box B' - components 2,3,4, 'Box C' - components 2,3.

$$\begin{aligned} P(\text{Box C operates}) &= P(\text{component 2 operates and component 3 operates}) \\ &= P(\text{component 2 operates})P(\text{component 3 operates}) = 0.4 \times 0.7 = 0.28. \end{aligned}$$

$$\begin{aligned} P(\text{Box B operates}) &= P(\text{Box C operates or component 4 operates}) \\ &= P(\text{Box C operates}) + P(\text{component 4 operates}) - \\ &\quad P(\text{Box C operates})P(\text{component 4 operates}) \\ &= 0.28 + 0.5 - 0.28 * 0.5 = 0.64. \end{aligned}$$

$$\begin{aligned} P(\text{system operates}) &= P(\text{component 1 and Box B operate}) \\ &= P(\text{component 1 operates})P(\text{Box B operates}) \\ &= 0.5 * 0.64 = 0.32. \end{aligned}$$

Thus, $P(\text{system does not operate}) = 0.68$

Q4. In a NiCd battery, a fully charged cell is composed of Nickelic Hydroxyde. Nickel is an element that has a multiple oxidation states. Let X be the nickel charge, which has the following probability mass function:

x	$f_X(x)$
0	.18
1	k
2	.33
4	.15

Determine the value of k and the mean of the nickel charge.

(a)	0.34, 2.0	(b)	0.15, 1.5	(c)	0.34, 1.6
(d)	1.45, 0.15	(e)	None of the preceding		

Q5. In the inspection of tin plate produced by a continuous electrolytic process, 1 imperfection is spotted per minute, on average. Find the probability of spotting at most two imperfections in 5 minutes. Assume that we can model the occurrences of imperfections as a Poisson process.

(a)	$\frac{3}{2}\exp(-1)$	(b)	$6\exp(-5)$	(c)	$\frac{37}{2}\exp(-5)$
(d)	$\exp(-1)$	(e)	None of the preceding		

Solution to Q5:

Rate per 5 minutes: $\lambda = 5$. If $X \sim Poisson, \lambda$, then to compute $P(X \leq 2) = P(X=0)+P(X=1)+P(X=2)$:

$$P(X=0)+P(X=1)+P(X=2) = \exp(-\lambda)+\exp(-\lambda)\lambda+\lambda^2/2\exp(-\lambda) = \frac{37}{2}\exp(-5)$$

Q6. Assume that X is a continuous random variable with the density f_X:

$$f_X(x) = \begin{cases} Ke^{-3x} & \text{if } x \geq -\ln(3)/3, \\ 0 & \text{otherwise.} \end{cases}$$

The value of K is:

(a) $\frac{\ln 3}{3}$ (b) $-\frac{\ln 3}{3}$ (c) 1

(d) 10 (e) None of the preceding.

Solution to Q6:

$$1 = \int_a^\infty f_X(x)\,dx = \int_{-\ln(3)/3}^\infty \exp(-3x)\,dx = \frac{1}{3}\exp(-3x)|_{x=-\ln(3)/3}^\infty = 1 \text{ so that } K = 1.$$

Q7. A random sample of size $n_1 = 16$ is selected from a normal population with a mean of 75 and variance of 288. A second random sample of size $n_2 = 9$ is taken independently from another normal population with mean 80 and variance of 162. Let $\bar{X}_1$ and $\bar{X}_2$ be the two sample means. Find the probability that $\bar{X}_1 + \bar{X}_2$ **is less than** 158.

(a) 0.691462 (b) 0.308538 (c) 0.738548

(d) 0.423501 (e) None of the preceding

Solution to Q7:

$\overline{X}_1 + \overline{X}_2 \sim N(75 + 80, 288/16 + 162/9) = N(155, 36)$.

$$P(\overline{X}_1 + \overline{X}_2 < 158) = P\left(Z < \frac{158 - 155}{\sqrt{36}}\right) = \Phi(0.5) = 0.691462.$$

Q8. Assume that we have a sample of size 10 from a population $N(4, 9)$. Denote by $\overline{X}$ and S^2, the sample mean and sample variance, respectively. Find c such that

$$P\left(\frac{\overline{X} - 4}{S/\sqrt{10}} \leq c\right) = .99.$$

(a) 1.833 (b) 2.326 (c) 1.645

(d) 2.821 (e) None of the preceding

Solution to Q8:

Equivalent statement: find c such that

$$P\left(\frac{\overline{X} - 4}{S/\sqrt{10}} \geq c\right) = 0.01.$$

We have that $\frac{\overline{X} - 4}{S/\sqrt{10}}$ has Student distribution with $n - 1 = 9$ degrees of freedom. From the table we read that $P(t_9 > 2.821) = 0.05$, thus $c = 2.821$.

Q9. In a random sample of 1000 houses in a certain city, it is determined that 228 are heated by oil. Find a 99% confidence interval for the number of houses in this city that are heated by oil.

(a) [0.202, 0.254] (b) [0.197, 0.259] (c) [0.194, 0.262]

(d) [0.185, 0.247] (e) None of the preceding

Q10. An article in *Computers and Electrical Engineering* considered the speed-up of cellular neural networks (CNN) for a parallel general-purpose computing architecture. The data are as follows:

3.77 , 3.35 , 4.21 , 4.03 , 4.03 , 4.63
4.63 , 4.13 , 4.39 , 4.84 , 4.26 , 4.60

Assume that the population is normally distributed. The 95% confidence interval for the mean speed-up is:

(a) [4.155,4.323] (b) [4.021,4.456] (c) [4.040,4.438]

(d) [3.77,4.60] (e) None of the preceding.

Solution to Q10:

```
> x=c(3.77 , 3.35 , 4.21 , 4.03 , 4.03 , 4.63,4.63 , 4.13 , 4.39 , 4.84 , 4.26 , 4.60)
> alpha=0.05
> n=length(x)
> mean(x)-qt(1-alpha/2,n-1)*sd(x)/sqrt(n)
[1] 3.972966
> mean(x)+qt(1-alpha/2,n-1)*sd(x)/sqrt(n)
[1] 4.505367
```

Answer E. It was not intentional

Q11. We want to test the hypothesis that the average content of containers of a particular lubricant equals 10 liters against the two-sided alternative. The contents of a random sample of 10 containers are

10.2	9.7	10.1	10.3	10.1
9.8	9.9	10.4	10.3	9.5

Find the P-value of this one-sided test. Assume that the distribution of contents is normal.

(a) $0.05 < P < 0.10$ (b) $0.10 < P < 0.20$ (c) $0.25 < P < 0.40$

(d) $0.50 < P < 0.80$ (e) None of the preceding

Solution to Q11:

We have $\sum_{i=1}^{n} x_i^2 = 1006.79$. If you use this, then: We test $H_0 : \mu = 10$ vs. $H_1 : \mu > 10$. We have $\overline{x} = 10.03$ and $s^2 = 0.08678$. The observed value of test statistics is

$$t_0 = \frac{\overline{x} - 10}{s/\sqrt{n}} = \frac{10.03 - 10}{\sqrt{0.08678}/\sqrt{10}} = 0.322$$

The P-value is $P(T > .322)$ and is between 0.25 and 0.40. Multiply by two.

For questions 12, 13:
An engineer measures the weight of pieces of steel. The weight follows normal distribution with variance 16.

Q12. The engineer measures $n = 25$ pieces of steel and obtains $\bar{x} = 6$. The two-sided 95% confidence interval for the mean is:

(a) (-0.272,12.272) (b) (4.432,7.568)

(c) (3.250,8.750) (d) (4.120,7.522)

(e) None of the preceding.

Solution to Q12:

$$(\bar{x} - z_{\alpha/2}\sigma/\sqrt{n}, \bar{x} + z_{\alpha/2}\sigma/\sqrt{n}) = (6 - 1.96 * 4/5, 6 + 1.96 * 4/5) = (4.432, 7.568)$$

Q13. The engineer measures $n = 25$ pieces of steel and obtains $\bar{x} = 6$. He wants to test $H_0 : \mu = 5$ against $H_1 : \mu > 5$. The p-value for the test is:

(a) 0.05000 (b) 0.10565 (c) 0.89435

(d) 1.0000 (e) None of the preceding.

Solution to Q13:

$$P(\bar{X} > 6) = P\left(Z > \frac{6-5}{4/5}\right) = P(Z > 1.25) = 1 - 0.89435 = 0.10565.$$

Q14. A group of 1000 adult American Catholics were asked "do you think there should be female priests?". 600 among them answered "yes". Let p be the true proportion of adult American Catholics that would allow female priests. Consider the test H_0: $p = 0.65$ against $H_1 : p < 0.65$. Determine the p-value and conclusion of the test at $\alpha = 0.05$. (p-value is rounded to 4th decimal place).

(a) 0.0005, reject H_0; (b) 0.0005, do not reject H_0;

(c) 0.9995, reject H_0; (d) 0.9995, do not reject H_0;

(e) None of the preceding.

Q15. The thickness of a plastic film (in mils) on a substrate material is thought to be influenced by the temperature at which the coating is applied. A completely randomized experiment is carried out. Eleven substrates are coated at 125^oF, resulting in a sample mean coating thickness of $\bar{x}_1 = 103.5$ and a sample standard deviation of $s_1 = 10.2$. Another 11 substrates are coated at 150^oF, for which $\bar{x}_2 = 99.7$ and $s_2 = 11.7$ are observed. We want to test equality of means against the two-sided alternative. The value of the appropriate test statistics and the decision are ($\alpha = 0.05$):

(a) 0.81; Reject H_0. (b) 0.81; Do not reject H_0. (c) 1.81; Reject H_0.

(d) 1.81; Do not reject H_0. (e) None of the preceding

Hint: You may consider that population variances are unknown but equal.

Solution to Q15:

This is two-sample test. $s_p^2 = \dfrac{10 * 10.2^2 + 10 * 11.7^2}{20} = 120.465$, $s_p = 10.97$. The observed value of the test statistics

$$\frac{\bar{x}_1 - \bar{x}_2}{s_p\sqrt{1/n_1 + 1/n_2}} = \frac{103.5 - 99.7}{10.97 * \sqrt{1/101 + 1/11}} = 0.81.$$

$t_{0.05/2,20} = 2.086$, don't reject H_0

For Questions 16, 17 use the data below:

Envoi	1	2	3	4	5	6	7	8	9	10
x	825	215	1070	550	480	920	1350	325	670	1215
y	3.5	1.0	4.0	2.0	1.0	3.0	4.5	1.5	3.0	5.0

Hint: $\sum_{i=1}^{n} x_i^2 = 7104300$, $\sum_{i=1}^{n} y_i^2 = 99.75$, $\sum_{i=1}^{n} x_i y_i = 26370$.

Q16. A company employs 10 part-time drivers for its fleet of trucks. Its manager wants to find a relationship between number of kilometers driven (X) and number of working days (Y) in a typical week. (The drivers are hired to drive half-day shifts, so that 3.5 stands for 7 half-day shifts). The manager wants to use the linear regression model

$$Y = \beta_0 + \beta_1 x + \epsilon.$$

The fitted regression line is:

(a) $\hat{y} = 0.1181 + 0.0036x$

(b) $\hat{y} = 117.15 - 0.15x$

(c) $\hat{y} = 5.59 - 0.0036x$

(d) $\hat{y} = 2.85 - 0.1521x$

(e) None of the preceding.

Solution to Q16:

$S_{xx} = \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} x_i\right)^2 = 1297860$, $S_{xy} = 4653$, so that $\hat{\beta}_1 = S_{xy}/S_{xx} = 0.0036$, $\hat{\beta}_1 = \sum_{i=1}^{n} y_i/n - \hat{\beta}_1 \sum_{i=1}^{n} x_i =$ 0.1181.

Q17. We want to test significance of regression, i.e. $H_0 : \beta_1 = 0$ against $H_1 : \beta_1 \neq 0$. The value of the appropriate statistics and the for $\alpha = 0.05$ decision is:

(a) 8.55; Do not reject H_0.

(b) 2.35; Reject H_0.

(c) 8.55; Reject H_0.

(d) 2.35; Do not reject H_0.

(e) None of the preceding.

Solution to Q17:

The estimated variance is

$$\hat{\sigma}^2 = \frac{S_{yy} - \hat{\beta}_1 S_{xy}}{n-2} = \frac{1.8434}{8} = 0.23.$$

Consequently,

$$t_0 = \frac{\hat{\beta}_1}{\sqrt{\hat{\sigma}^2/S_{xx}}} = \frac{0.0036}{\sqrt{0.23/1297860}} = 8.551701.$$

Now, $t_{0.025,8} = 2.306$ - reject H_0.

Q18. Ten individuals have participated in a diet modification program to stimulate weight loss. Their weight both before and after participation in the program is shown below:

Before	195, 213, 247, 201, 187, 210, 215, 246, 294, 310
After	187, 195, 221, 190, 175, 197, 199, 221, 278, 285

Is there evidence to support the claim that this particular diet-modification program is effective in producing mean weight reduction? Compute P-value and state your conclusion using $\alpha = 0.05$.

(a) p-value< 0.0005; reject H_0

(b) P-value< 0.0005; do not reject H_0

(c) $0.25 < p$-value< 0.40; reject H_0

(d) $0.25 < p$-value< 0.40; do not reject H_0

(e) None of the preceding

Solution to Q18:

This is the paired t-test. Compute difference (After-Before):

Before	195, 213, 247, 201, 187, 210, 215, 246, 294, 310
After	187, 195, 221, 190, 175, 197, 199, 221, 278, 285
D_i	$-8, -18, -26, -11, -12, -13, -16, -25, -16, -25$

To test $H_0 : \mu_D = 0$ vs. $H_0 : \mu_d < 0$. We have $\bar{d} = -17$, $s_D^2 = 41.11$. The observed value of test statistics is

$$t_0 = \frac{\bar{d}}{s_D/\sqrt{10}} = -8.38.$$

The P-value is

$$P(t_9 < -8.38) = P(t_9 > 8.38) < 0.0005,$$

reject H_0 - evidence that the diet reduces a weight.

For Questions 19, 20 use the data below:
We have $m = 5$ preliminary samples of size $n = 3$ (some numbers are missing on purpose):

i				$\bar{x}_i$	r_i	s_i
1	27.1	29.4		27.9		1.3
2	30.6	32.5	32.4	31.83	1.9	1.07
3	25.7	35.5	30	30.4		4.91
4	31.1	23.2	25	26.43	7.9	
5	24.1	34.2	27.4	28.57	10.1	5.15
		total		145.13	32	16.57

Q19. The Control Chart for $\bar{X}$ from $\bar{R}$ is:

(a) [22.4788,35.5732]

(b) [22.5490,35.5031]

(c) [29.026,35.5031]

(d) [22.4134,36.2217]

(e) none of the preceding

Solution to Q19:

The estimated grand mean is $145.13/5 = 29.026$; $\bar{r} = 32/5 = 6.4$. The Control Chart for $\bar{X}$ from $\bar{R}$ is

$$\begin{aligned} CL &= \hat{\mu} = \bar{\bar{x}} = 29.026 \\ UCL &= \bar{\bar{x}} + A_2\bar{r} = 29.026 + 1.023 \times 6.4 = 35.5732 \\ LCL &= \bar{\bar{x}} - A_2\bar{r} = 22.4788. \end{aligned}$$

Q20. The Control Chart for $\bar{X}$ from $\bar{S}$ is:

(a) [22.4788,35.5732] (b) [29.026,35.5031] (c) [22.4134,36.2217]

(d) [22.5490,35.5031] (e) none of the preceding

Solution to Q20:

The estimated grand mean is $145.13/5 = 29.026$; $\bar{s} = 3.314$. The Control Chart for $\bar{X}$ from $\bar{S}$ is

$$\begin{aligned} CL &= \hat{\mu} = \bar{\bar{x}} = 29.026 \\ UCL &= \bar{\bar{x}} + \frac{3}{c_4\sqrt{n}}\bar{s} = 29.026 + \frac{3}{0.8862 \times \sqrt{3}} \times 3.314 = 35.5031 \\ LCL &= \bar{\bar{x}} - \frac{3}{c_4\sqrt{n}}\bar{s} = 29.026 - \frac{3}{0.8862 \times \sqrt{3}} \times 3.314 = 22.5490. \end{aligned}$$

Q21. The following output was produced with `t.test` command in `R`

```
One Sample t-test

data:  x

t = 2.0128,df = 99, p-value = 0.02342

alternative hypothesis: true mean is greater than 0
```

Based on this output, which statement is correct:

(a) If the type I error is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;

(b) If the type I error is 0.05, then we reject $H_0 : \mu = 0$ in favour of $H_1 :\neq 0$;

(c) If the type I error is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu > 0$;

(d) If the type I error is 0.01, then we reject $H_0 : \mu = 0$ in favour of $H_1 : \mu < 0$;

(e) Type I error is 0.02342.

Q22. Consider the following R output:

```
> pbinom(15,100,0.25)
[1] 0.01108327
> pbinom(16,100,0.25)
[1] 0.02111062
> pbinom(17,100,0.25)
[1] 0.03762626
> pbinom(30,100,0.25)
[1] 0.8962128
> pbinom(31,100,0.25)
[1] 0.9306511
> pbinom(32,100,0.25)
[1] 0.9554037
```

Let X be a binomial random variable with $n = 100$ and $p = 0.25$. Using the R output above, calculate $P(16 \leq X < 31)$.

(a)	0.9196	(b)	0.9095	(c)	0.9348
(d)	0.8851	(e)	none of the preceding		

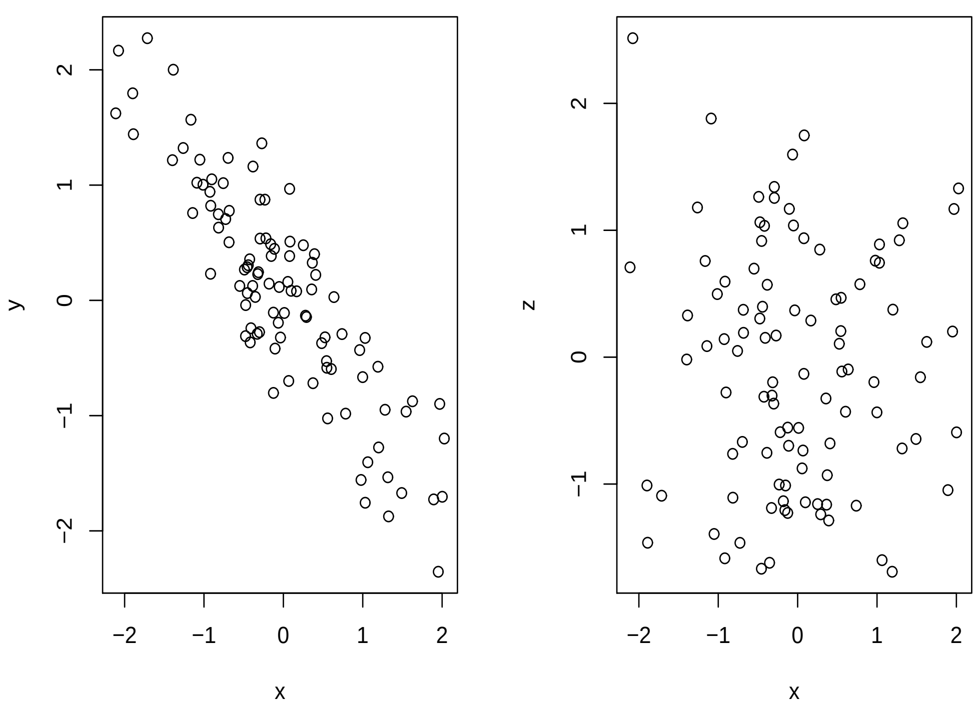


Figure 1. Figure for Question 23

Q23. For bivariate data displayed on Figure 1, the correlation coefficients are, respectively:

(a) 0.9, 0; (b) -0.9, 0; (c) 0.1, 0.9;

(d) 0.5, 0.5; (e) none of the preceding.

Q24. A pharmaceutical company claims that a drug decreases a blood pressure. A physician doubts this claim. He tests 10 patients and records results before and after the drug treatment:

```
Before=c(140,135,122,150,126,138,141,155,128,130)

After=c(135,136,120,148,122,136,140,153,120,128)
```

He is a big fan of `R`, so that he types

```
test.t(Before,After,alternative="greater")
```

and obtains

```
data:  Before and After

t = 0.5499, p-value = 0.2946

alternative hypothesis: true difference in means is greater than 0

sample estimates: mean of x mean of y
    136.5     133.8
```

His assistant claims that the command should be

```
test.t(Before,After,paired=TRUE,alternative="greater")
```

He obtains

```
data: Before and After t = 3.4825, df = 9, p-value = 0.003456
alternative hypothesis: true difference in means is greater than 0
sample estimates: mean of the differences
                    2.7
```

(a) The assistant uses the correct command. There is not enough evidence to justify that the new drug decreases blood pressure;

(b) The assistant uses the correct command. There is enough evidence to justify that the new drug decreases blood pressure for any reasonable choice of α;

(c) The physician uses the correct command. There is not enough evidence to justify that the new drug decreases blood pressure;

(d) The physician uses the correct command. There is enough evidence to justify that the new drug decreases blood pressure for any reasonable choice of α;

(e) Nobody is correct, t-test should not be used here.

This is the last question

Solutions to multiple choice questions:

Q1 $\longrightarrow$ d

Q2 $\longrightarrow$ d

Q3 $\longrightarrow$ d

Q4 $\longrightarrow$ c

Q5 $\longrightarrow$ c

Q6 $\longrightarrow$ c

Q7 $\longrightarrow$ a

Q8 $\longrightarrow$ d

Q9 $\longrightarrow$ c

Q10 $\longrightarrow$ e

Q11 $\longrightarrow$ d

Q12 $\longrightarrow$ b

Q13 $\longrightarrow$ b

Q14 $\longrightarrow$ a

Q15 $\longrightarrow$ b

Q16 $\longrightarrow$ a

Q17 $\longrightarrow$ c

Q18 $\longrightarrow$ a

Q19 $\longrightarrow$ a

Q20 $\longrightarrow$ d

Q21 $\longrightarrow$ a

Q22 $\longrightarrow$ d

Q23 $\longrightarrow$ b

Q24 $\longrightarrow$ b