Part I. **(25 points)** In Supradyn All Day Vitamin advertisement, the following statistics is announced by Burcu Esmersoy: 77% of people in Turkey press the snooze button (erteleme düğmesi) of their alarm clock at least once every morning.

1. (~~5~~ 3 points) In a random sample of 9 subjects who use an alarm clock, what is the probability that exactly 8 of them press the snooze button?

$$\binom{9}{8}(0.77)^8(0.33)^1 = 9\,(0.123)(0.23) = 0.254$$

2. (~~8~~ 7 points) What is the probability that at most 2 of them press the snooze button in a random sample of 9?

$$\binom{9}{0}(0.77)^0(0.33)^9 + \binom{9}{1}(0.77)^1(0.33)^8 + \binom{9}{2}(0.77)^2(0.33)^7$$

$$= (4.64)10^{-5} + (9.75)10^{-4} + (9.1)10^{-3} \cong 0.01$$

3. (6 points) If it is found in a random sample of 9 subjects that 2 of them pressed the snooze button at least once every morning, is the claim of the advertisement supported by this evidence or not? Hint: Use your answer in part 2.

$H_0: p = 0.77$

$H_a: p < 0.77$

P-value is about 1%, small.

So we reject H_0.

It is more likely that true percentage is less than 77%.

4. (9 points) If it is found in a random sample of 99 subjects that 68 of them press the snooze button at least once every morning, is the claim of the advertisement supported or not?

$H_0: p = 77\%$

$H_a: p < 77\%$

$$\hat{p} = \frac{68}{99} \approx 0.69$$

$$SE = \sqrt{\frac{(0.77)(0.23)}{99}} \cong 0.042$$

$$z = \frac{0.69 - 0.77}{0.042} = -1.9$$

$$\text{P-value} = \frac{100\% - 94.26\%}{2} \cong 2.87\% < 5\%$$

$\Rightarrow$ We reject H_0.

The claim of 77% is not supported.

Part II. (15 points) A self-esteem questionnaire is used to collect data from two groups, those who left their job within a few months after graduation (leavers) and those who remained in their job after they graduated (stayers). The summary statistics are as follows for the respective self-esteem scores.

	Leavers	Stayers
Mean	3.05	2.96
SD	0.8	0.7
Sample Size	103	225

1. (9 points) At $\alpha = 0.01$, can it be concluded that leavers have a higher self-esteem than stayers?

$H_0: \mu_1 = \mu_2$

$H_a: \mu_1 > \mu_2$

$$SE_{diff} = \sqrt{\frac{(0.8)^2}{103} + \frac{(0.7)^2}{225}} \cong 0.09$$

$$z = \frac{3.05 - 2.96}{0.09} = 1$$

$$p.value = \frac{100 - 68.27}{2}\% = 15.865\% > 1\%$$

So we do not reject H_0.

There is no sufficient evidence that leavers have a higher self-esteem than stayers.

$\rightarrow z \cong 1.65$

2. (6 points) Construct a 90% confidence interval for the self-esteem score of stayers.

$$2.96 \mp 1.65 \frac{0.7}{\sqrt{225}}$$

$$2.96 \mp \underbrace{(1.65)(0.0467)}_{0.077}$$

$$(2.88, 3.04)$$

Part III. (20 points) A random sample of 150 beer drinkers were surveyed about their preference for three types of beers, namely light, regular, or dark.

	Light	Regular	Dark	
Male	20	40	20	80
Female	30	30	10	70
	50	70	30	150

1. (3 points) Define the variable(s) in this study.

Preference for beer type

Gender

2. (2 points) Identify the sample(s) in this study.

A single sample of 150 beer drinkers.

3. (10 points) Is beer preference independent of the gender of the beer drinker?

Expected frequencies →

26.7	37.3	16
23.3	32.6	14

$\frac{50 \cdot 80}{150} \cong 26.7$ etc

H_0: Gender and preference are independent

H_1: " " " " not independent

$d.f. = (2-1)(3-1) = 2$ $\quad \chi^2_{\text{critical}} = 5.99$ at $\alpha = 5\%$

$\chi^2 = \frac{(20-26.7)^2}{26.7} + \dots + \frac{(10-14)^2}{14} = 6.13$

Since $6.13 > 5.99$, reject H_0. "The preference for beer type depends on gender."

4. (5 points) Find the conditional distribution of beer preference of females.

	Light	Regular	Dark
Females	42.9%	42.9%	14.2%

change just to make the total 100%

$\frac{30}{70} \cong 0.429$ $\quad \frac{10}{70} \cong 0.143$

Part IV. (15 points)

1. (8 points) Consider the sample given in Part III, pg.3. Find a 95% confidence interval for the difference of the percentages of female and male beer drinkers.

$$\hat{p}_M = \frac{80}{150} \cong 0{,}53 \qquad \hat{p}_F = \frac{70}{150} \cong 0{,}47$$

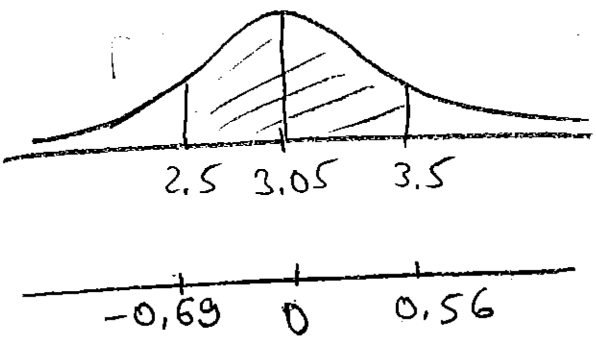
$$SE_{diff} = \sqrt{\frac{(0.53)(0.47)}{80} + \frac{(0.46)(0.54)}{70}} = 0.08$$

$$0{,}53 - 0{,}47 \mp 1{,}96\,(0{,}08)$$

$$(-0{,}09\ ,\ 0{,}22)$$

2. (7 points) Consider the self-esteem questionnaire in Part II, pg.2. Assume that the self-esteem scores have a normal distribution in the population, and estimate the population parameters with the given statistics. Then, approximately what proportion in the population of leavers has a self-esteem score between 2.5 and 3.5?

mean is estimated as 3.05
SD 〃 〃 〃 0,8



$$\frac{2{,}5 - 3{,}05}{0{,}8} \cong -0{,}69$$

$$\frac{3{,}5 - 3{,}05}{0{,}8} \cong 0{,}56$$

$$\text{Table} \Rightarrow \frac{41{,}77\%}{2} + \frac{51{,}61\%}{2} = 46.7\%$$

Part V. **(20 points)** In a large city, the typical price paid for a meal in a given restaurant is investigated by a newspaper reporter. The reporter randomly selected 7 Italian restaurants, 6 Seafood restaurants, and 8 Steakhouses and recorded the following prices.

	Italian	Seafood	Steakhouse
	$12	$16	$24
	13	18	19
	15	17	23
	17	26	25
	18	23	21
	20	15	22
	17		27
			31
Mean	16	19	24
Standard Deviation	2.8	4.6	3.7

$n = 21$

$$X_{GM} = \frac{16\times7 + 19\times6 + 24\times8}{21} = 19.9$$

1. (~~2~~ 3 points) Identify the variable(s), and the population(s) under investigation.

Variable: Price paid for a meal

Populations: 1) Italian restaurants 2) Seafood Restaurants 3) Steakhouses in the city, all of them!

2. (~~6~~ 8 points) Fill in the missing entries in the following table.

$k = 3$

$n = 7 + 6 + 8 = 21$

$n - k = 18$

$\boxed{k - 1 = 2}$

	Sum of Squares	d.f.	Mean square	F
Between	245.8	2	122.9	8.9
Within	248.67	18	13.8	

$$SS_{Between} = 7(16-19.9)^2 + 6(19-19.9)^2 + 8(24-19.9)^2$$

$$= 245.8$$

$$s_B^2 = \frac{245.8}{2} = 122.9 \quad s_W^2 = \frac{248.67}{18} \cong 13.8 \Rightarrow F = \frac{122.9}{13.8} = 8.9$$

3. (6 points) Can one conclude that there is a significant difference among the meal price for the three types of restaurants? $H_0: \mu_1 = \mu_2 = \mu_3$ H_a: at least one μ is different

$F_{2,18,0.05} \cong 3.55$, Since $8.9 > 3.55$, we reject H_0.

There is a significant difference between the average meal prices of the three types of restaurant

4. (3 points) State your assumption(s) for the test procedure you used in Question 3 to be applicable.

1) We assume that prices in each population are normally distributed.
2) The populations are independent.
3) The variances of the populations are equal.

Part VI. **(15 points)** A study for finding out how well airline companies serve their customers showed the following customer ratings: 3% excellent, 28% good, 45% fair, and 24% poor. In a later study of service by mobile phone companies, a sample of 400 adults indicated the following customer ratings: 24 excellent, 124 good, 172 fair, and 80 poor. Is the distribution of the customer ratings for phone companies different from the distribution of customer ratings for airline companies?

Expected frequencies:

Excellent	Good	Fair	Poor
12	112	180	96

3% of 400 : $\frac{3}{100} \cdot 400 = 12$ $\qquad \frac{45}{100} \cdot 400 = 180$

$\frac{28}{100} \cdot 400 = 112$ $\qquad \frac{24}{100} \cdot 400 = 96$

$$\chi^2 = \frac{(12-24)^2}{12} + \frac{(112-124)^2}{112} + \frac{(180-172)^2}{180} + \frac{(80-96)^2}{96}$$

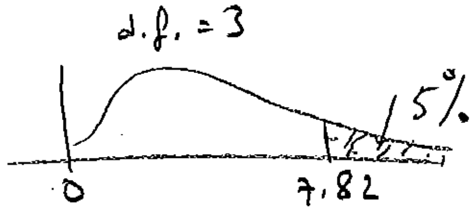
$$= 12 + 1{,}286 + 0{,}356 + 2{,}67$$

$$= 16{,}312$$

$\boxed{\text{d.f. } 4-1=3}$

H_0 : $p_1 = 3\%$, $p_2 = 28\%$, $p_3 = 45\%$, $p_4 = 24\%$.

H_a : at least one equality does not hold.



$$\chi^2_{0.05,\,3} = 7{,}82$$

Since $16{,}312 > 7{,}82$, we reject H_0.

So, the distribution of customer ratings for phone companies is significantly different from those of airline companies.