

NYU Tandon School of Engineering
MA-UY 2224 DATA ANALYSIS FINAL (A) MAY 14, 2018

Print Name:

Signature:

Section:

Instructor:

ID #:

Directions: You have **110 minutes** to answer the following questions. ***You must show all your work*** as neatly and clearly as possible and indicate the final answer clearly. You may use only a TI-30 calculator.

If you are feeling ill you should inform the proctor. The proctor will note your name, student ID and accept any written statement(s) that you may wish to make regarding your illness. Cell phones and other electronic devices may **NOT** be used during the exam.

Problem	Possible	Points
1	28	
2	18	
3	16	
4	16	
5	12	
6	10	
Total	100	

Mathematics Department After-Test Survey:

“We care what you think!”

1. This exam was

(a) too hard

(b) fair

(c) easy

2. This exam reflected the assigned homework/worksheet problems

(a) a lot

(b) somewhat

(c) not much

3. The time I was given to complete this exam was

(a) too much

(b) just right

(c) too little

(1) Fill in the blanks. **Show your work.**

- (a) The number of flaws on a certain type of aluminum screen follows a Poisson distribution, with an average of two flaws per 100 feet. Then the probability that there are less than 3 flaws on a **200-foot screen** is _____.

- (b) Suppose that the length of an order for a Halal food truck in minutes is an exponential random variable with probability density function (p.d.f.) $f(x) = \frac{1}{3}e^{-x/3}$, $x > 0$. If someone arrives immediately ahead of you at a food truck, the probability that you will have to wait between 3 and 6 minutes is _____.

-
- (c) According to a recent survey, about 60% of Americans do not exercise regularly. Randomly select 100 Americans, estimate the probability that between 50 and 65 of them, inclusive, do not exercise regularly.
- (d) Last semester, I taught two sections of a same class; Section A with 20 students and Section B with 30. Before grading their final exams, I randomly mixed all the exams together. I graded 12 exams at the first sitting.
- (i) Of those 12 exams, the probability that exactly 5 of these are from the Section B is _____. (You do not need to simplify your answers.)
- (ii) Of those 12 exams, the probability that they are not all from the same section is _____. (You do not need to simplify your answers.)

-
- (2) In the country of StatLand, the annual salary (in thousands of dollars) of all workers with a college degree has the distribution $N(68, 10^2)$, and of all workers without a college degree, $N(40, 8^2)$. Show your work.
- (a) Randomly select one worker with a college degree and two workers without a college degree. Find the probability that the worker with a college degree makes more than the combined salary of the two without the college degree.
- (b) Randomly select one worker from each of the two groups. Find the probability that the worker with a college degree makes at least \$25,000 more than the one without the college degree.
- (c) Randomly select 100 workers from each of the two groups. Find the probability that the average salary of 100 college graduates is at least \$25,000 higher than the average of the 100 without a college degree.

(3) A computer is tested by running a collection of benchmark programs and recording the CPU times (in seconds) required to complete the programs. A random sample of 6 benchmark programs produced a sample mean $\bar{x} = 15.53$ and sample standard deviation $s = 4.2$. Assume the running time is normally distributed.

(a) Test $H_0 : \sigma^2 = 9$ vs $H_1 : \sigma^2 \neq 9$ at significance level $\alpha = 0.10$. Make sure to show your work and state your conclusion clearly. Please use the approach of the critical region.

(b) Test $H_0 : \mu = 18$ vs $H_1 : \mu < 18$ at significance level $\alpha = 0.05$. Make sure to show your work and state your conclusion clearly. Please indicate the approximate p -value of this test.

(4) Show your work.

- (a) To study a freshmen class' performance in the midterm exam and final, the grades of 7 randomly selected students on a midterm exam (x) and on the final (y) are recorded as follows:

Student	#1	#2	#3	#4	#5	#6	#7
x	77	50	71	72	61	94	96
y	82	66	78	65	75	85	90

Assume both grades distributions are approximately normal, construct a 90% confidence interval for the difference in the mean scores between the midterm and the final.

- (b) A company held a raffle during an event. The designer of that raffle claimed that 1% of ticket holders would be awarded the grand prize, 5% would be awarded runner up prize, 10% would be awarded the third prize, and the rest would get no prize. Of the 2000 people who got a raffle ticket, 30 won the grand prize, 120 won the second prize, 210 won the third prize, and the rest got nothing. Use a chi-square goodness-of-fit test to test if the model the designer described is a good fit. Use the significance level $\alpha = 0.05$.

- (5) Let p equal the proportion of residents in New York City who believe that housing is too expensive. Suppose 1131 out of a random sample of 1450 residents from New York City believe so. Show your work when solving the following problems.

(a) Find a 95% confidence interval for p .

(b) How large a sample is required if we want to be 95% confident that our estimate of p in part (a) is within 0.01 of the true value?

(6) A biologist would like to test to see if the width of a 5-week-old dog's paw can be used to predict the height of the dog when it is fully grown. The biologist notes the width of the paws of 8 different puppies. When they are fully grown, their heights are measured. The results are found on the following page.

(a) If you were going to fit the data with a least square regression line, what is the equation of the line?

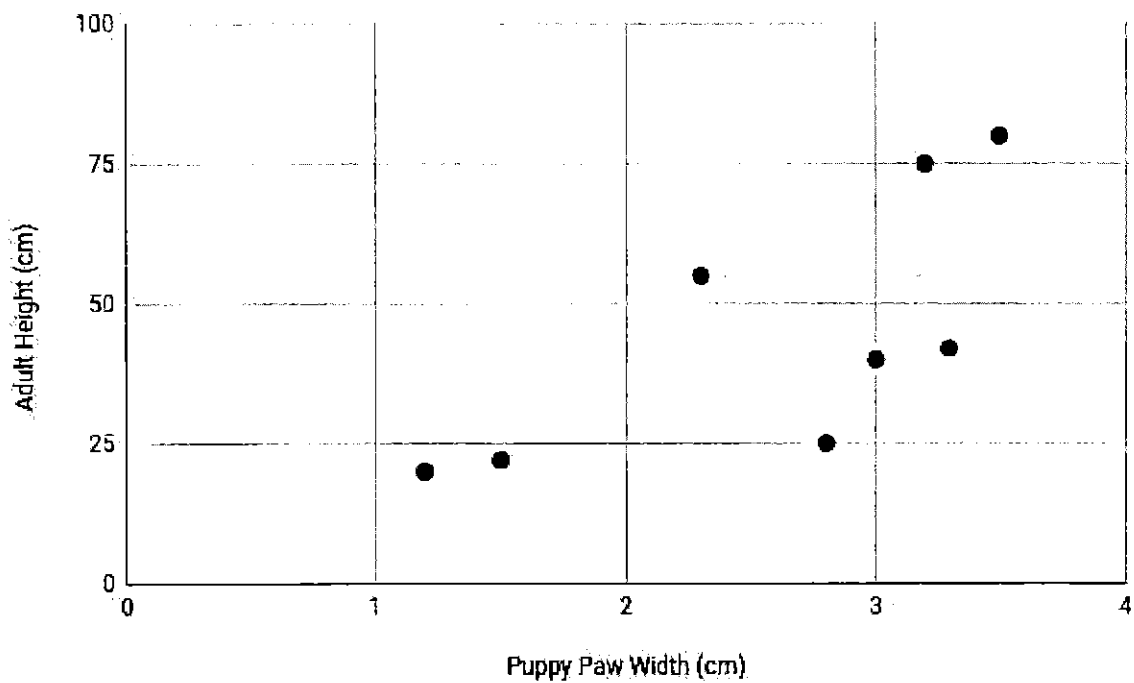
(b) What is the proportion of variation in the heights of the dogs that can be explained by using the widths of their paws as puppies as a predictor? In other words, find the coefficient of determination.

(c) Test $H_0 : \beta_1 = 0$ vs. $H_1 : \beta_1 \neq 0$ at significance level $\alpha = 0.05$. Do you think the width of the paw of a puppy can be used to predict its adult height?

(d) Find a 95% prediction interval of the eventual height of a dog that is currently 5 weeks old and has a paw of width 4 cm.

Puppy Paw Width (cm)	Adult Height (cm)
1.2	20
2.3	55
3.5	80
2.8	25
3.2	75
1.5	22
3	40
3.3	42

\bar{x}	2.6
\bar{y}	44.875
S_{xx}	5.12
S_{xy}	98.7
S_{yy}	3812.875



Puppy Paw Width (cm)	Adult Height (cm)
1.2	20
2.3	55
3.5	80
2.8	25
3.2	75
1.5	22
3	40
3.3	42

\bar{x}	2.6
\bar{y}	44.875
S_{xx}	5.12
S_{xy}	98.7
S_{yy}	3812.875

