UNIVERSITY OF BRITISH COLUMBIA
DEPARTMENT OF MECHANICAL ENGINEERING
FINAL EXAMINATION, APRIL 2018

MECH 305/306 – Data Analysis and Mech Labs

Duration: Target = 2 hours. Maximum = 2.5 hours. Answer all 5 questions.

Materials admitted: Non-communicating, non-programmed calculator, personal handwritten notes within 2 letter sheets, both sides.

The purpose of this exam is to evaluate your knowledge of the course material. Orderly presentation demonstrates your knowledge most clearly, while disorganized and unprofessional work creates serious doubt. Marks are assigned accordingly. A bonus of up to 5 marks will be given for exemplary presentation.

Write your name on *each* page *during* the examination time.

NAME: ______________________ SIGNATURE: ______________________________

SECTION: ___________________ STUDENT NUMBER: _________________________

Student Conduct During Examinations

1. Each examination candidate must be prepared to produce, upon the request of the invigilator or examiner, his or her UBCcard for identification.
2. Examination candidates are not permitted to ask questions of the examiners or invigilators, except in cases of supposed errors or ambiguities in examination questions, illegible or missing material, or the like.
3. No examination candidate shall be permitted to enter the examination room after the expiration of one-half hour from the scheduled starting time, or to leave during the first half hour of the examination. Should the examination run forty-five (45) minutes or less, no examination candidate shall be permitted to enter the examination room once the examination has begun.
4. Examination candidates must conduct themselves honestly and in accordance with established rules for a given examination, which will be articulated by the examiner or invigilator prior to the examination commencing. Should dishonest behaviour be observed by the examiner(s) or invigilator(s), pleas of accident or forgetfulness shall not be received.
5. Examination candidates suspected of any of the following, or any other similar practices, may be immediately dismissed from the examination by the examiner/invigilator, and may be subject to disciplinary action:
6. i. speaking or communicating with other examination candidates, unless otherwise authorized;
   ii. purposely exposing written papers to the view of other examination candidates or imaging devices;
   iii. purposely viewing the written papers of other examination candidates;
   iv. using or having visible at the place of writing any books, papers or other memory aid devices other than those authorized by the examiner(s); and,
   v. using or operating electronic devices including but not limited to telephones, calculators, computers, or similar devices other than those authorized by the examiner(s) — (electronic devices other than those authorized by the examiner(s) must be completely powered down if present at the place of writing).
7. Examination candidates must not destroy or damage any examination material, must hand in all examination papers, and must not take any examination material from the examination room without permission of the examiner or invigilator.
8. Notwithstanding the above, for any mode of examination that does not fall into the traditional, paper-based method, examination candidates shall adhere to any special rules for conduct as established and articulated by the examiner.
9. Examination candidates must follow any additional examination rules or directions communicated by the examiner(s) or invigilator(s).

CANDIDATES MUST IMMEDIATELY STOP WRITING WHEN THE INVIGILATOR ANNOUNCES THE EXAM IS OVER.

#	Mark	Max.	#	Mark	Max.
1		20	5		20
2		20	Pres.		5
3		20	**Total**		**100+5**
4		20			

Name ______________________ Student Number ______________

1. Circle the best answer for questions a,b,c,d,e, and g; give a short explanation in questions f) and h).

a) Flaws (defects) on a fiber optic cable occur randomly. That is, the location of the next flaw is independent of the previous flaws. The distance x between flaws is most likely distributed by

   i) binomial distribution
   ii) Poisson distribution
   iii) exponential distribution
   iv) Chi-square distribution
   v) normal distribution
   vi) log-normal distribution

Classic case of a "Poisson process", but the Poisson distribution provides the number of flaws per length, where as the spacing is a exponential. Poisson and exponential distributions are like Siamese twins.

b) Airborne particulate matter can range in size from a few nanometers (0.001 microns) to 10 microns, with a median diameter of 0.1 microns. 2/3 of the particles are between 0.05 microns and 0.2 microns. The pdf of particle size is most likely to follow which distribution?

   i) binomial distribution
   ii) Poisson distribution
   iii) exponential distribution
   iv) Chi-square distribution
   v) normal distribution
   vi) log-normal distribution

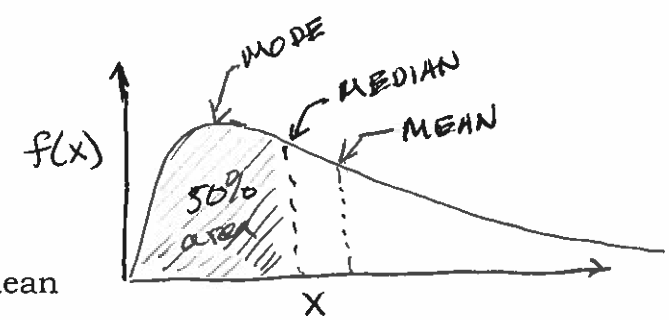
.05 is half of the median, and .2 is twice the median, so if you take logs, these points are equidistant from the median. Given that 2/3 of particles are within this range, we infer that this is really +/- one standard deviation of the median - if we have a normal distribution IN LOG SPACE.

c) The probability of a student getting a grade above 90% in Mech 305/306 is 0.15. The number of students out of a class of 130 getting above 90% is a random variable described by which of the following distributions?:

   i) binomial distribution
   ii) Poisson distribution
   iii) exponential distribution
   iv) Chi-square distribution
   v) normal distribution
   vi) log-normal distribution

d) The median of a log-normal distribution:

   i) is equal to the mean
   ii) is greater than the mean
   iii) is less than the mean
   iv) could be greater or less than the mean
   v) could be greater, equal or less than the mean



Name ______________________________ Student Number ______________

e) Five different experiments were done to compare pairs of samples (all called A or B, but they are different quantities in each experiment). The means, number of measurements in each sample, and the standard deviations are given below. Which experiment exhibits the most significant difference between A and B? ***(circle the experiment number on the table).***

Experiment	Sample	N	mean	standard deviation	SDOM $\sigma/\sqrt{N}$	δ	δ/SDOM
1	A	10	6	0.5	.158	0.2	1.3
	B	10	6.2	0.5			
2	A	20	8	1	.224	1.0	4.5
	B	20	7	1			
3	A	10	12	0.3	.095	2.0	21.1
	B	10	14	0.3			
4	A	100	13	0.2	.020	0.1	5.0
	B	100	13.1	0.2			
5	A	1000	51	3	.095	0.5	5.3
	B	1000	51.5	3			

You might notice that when computing the difference , the standard deviation should increase by sqrt(2) by error propagation rules. However this applies equally to all experiments, so the relative numbers are fine

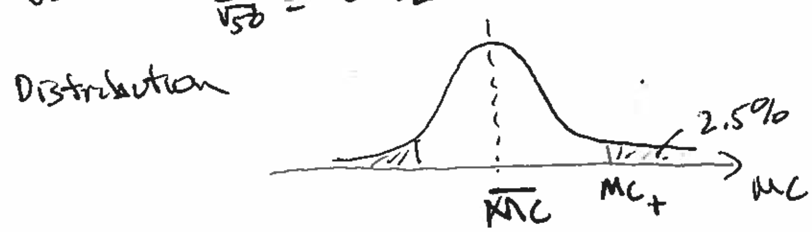
f) For question e), what combination of variables did you use to assess the significance of the difference?

Compare δ with uncertainty in means , $(\delta/\sigma)\sqrt{N}$
because this will be used in Z or t- test.

g) A sawmill produces kiln-dried dimension lumber with a target Moisture Content MC = 12. It is found from experience that the typical standard deviation of the moisture contents within the boards is MC = 0.9. What critical region for the average MC value of a 50-board sample would you set up to assess whether the MC = 12 target value is being reached with a p-value of 5% ? *(Hint: you may use the table on p.15.)*

i) 11.79 to 12.21
ii) 10.5 to 13.5
iii) 12.0 to 12.21
iv) 10.2 to 13.8
v) 11.75 to 12.25
vi) 12.0 to 13.5

SDOM = $\frac{0.9}{\sqrt{50}}$ = 0.1273



h) For question g), what statistic did you use for the assessment and how was this computed from the given information?

Assume 2-tailed test, tail area = 2.5% or Z ≈ 1.95
Therefore, the half-width of critical region is 1.95 (.1273)
= 0.25
Correct critical region is 11.75 to 12.25,

Setting this up as 1-tail test is also possible, 12.0 - 12.21.
(the 1 tail version seems less natural).

Name ______________________ Student Number ______________

2. Interpretation of measurements is greatly aided by an understanding of the measurement errors involved. Suppose that we wish to determine the Young's Modulus E for a cantilever beam by measuring the deflection caused by load at the free end. The relationship between deflection and load is:

$$\delta = \frac{F L^3}{3 EI} \qquad \text{where} \qquad I = \frac{b h^3}{12}$$

a) Design a test plan for evaluating the Young's Modulus. Prepare a blank table with labeled columns that you would use for recording your data. Explain your analysis method.

b) For a 1% variation within each of the measured quantities, determine the corresponding variations in computed E value.

c) Identify the aspects of your proposed measurements that may be expected to be the major contributors to the overall measurement error. Describe the precautions you would take to minimize the effect of the major error sources.

d) You would like to determine 95% confidence intervals on the value of E. You could do this using repeated trials, or by considering error propagation of the component measurements. Which method would you choose, and why?

Key ideas (not full solution)

a) $E = \frac{4 F L^3}{\delta b h^3}$ measured quantities are F, L, δ, b, h

b) Error propagation will show E most sensitive to L, h.

c) • L & h both critical, but h is smaller and may not be uniform. controlling this is critical.

• δ is likely very small, so a dial gauge or similar instrument is needed.

• Setting correct clamped boundary condition is also important.

• other disussion is possible and was accepted.

d) Repeated trials, and ideally calibration with a beam of known E.

Of course, most students provided more details (which was good!) but I provide here only the bare bones of a passing answer. The reason I favor doing some repeated trials and calibration is that error propagation is only as good as the uncertainties that you put into the propagation. More often than not, these are not well known. Error propagation is also only as good as the model of the experiment, including all factors. There may be additional factors such as temperature that affect the results.

Name ______________________ Student Number ______________

3. Answer each question and explain your reasoning.

(a) There are eight empty parking spaces in a row on one side of a city street. Two cars come by and park in randomly chosen spaces. On average, what is the number of empty parking spaces between them?

There are eight equally probable parking places for the first car. Because of left-to-right symmetry, we need consider only the first four. Let "F" be the position of the first car and the numbers equal the number of spaces between the second and first car.

F	0	1	2	3	4	5	6
0	F	0	1	2	3	4	5
1	0	F	0	1	2	3	4
2	1	0	F	0	1	2	3

0 spaces occurs 7 times.
1 spaces occurs 6 times. etc

Average $= 7\times0 + 6\times1 + 5\times2 + 4\times3 + 3\times4 + 2\times5 + 1\times6 / 28$

$= 56/28 = \underline{2}$

(b) Five students with different birthdays go out for dinner at a Chinese restaurant and sit at a circular table with five seats. They then notice that they have ended up sitting in order of their ages (not including the cyclic jump back from the oldest to the youngest). What is the probability of this happening?

The first person can sit anywhere. → Probability = 1

There are 4 seats available for the second person.

The sequence could be CW or CCW → a possible position on either side. → Probability $= \frac{2}{4}$

There are 3 seats available for third person → prob $= \frac{1}{3}$

There are 2 seats available for fowth person → prob $= \frac{1}{2}$

Fifth person has no choice → prob. = 1

Overall probability $= 1 \times \frac{2}{4} \times \frac{1}{3} \times \frac{1}{2} \times 1 = \underline{\frac{1}{12}}$

Name ______________________ Student Number ______________

(c) In a gambling game, the player pays \$1 to enter, names a number between 1 and 6 and then throws three dice. The player wins \$1 if the named number comes up once, \$3 on two dice and \$5 on all three dice. On average, how much the player win or lose per game?

For each die, the probability of getting the named number is $\frac{1}{6}$ and of not getting it is $\frac{5}{6}$. For three dice, there are 8 binary combination, 0 = not number, 1 = number

combination	probability × 216
0 0 0	$5 \times 5 \times 5 = 125$
0 0 1	$5 \times 5 \times 1 = 25$
0 1 0	$5 \times 1 \times 5 = 25$
0 1 1	$5 \times 1 \times 1 = 5$
1 0 0	$1 \times 5 \times 5 = 25$
1 0 1	$1 \times 5 \times 1 = 5$
1 1 0	$1 \times 1 \times 5 = 5$
1 1 1	$1 \times 1 \times 1 = 1$

Cost = \$1

Error here: should be 125/216 not 175/216

Avg. payback

$$= \frac{125 \times 0 + 3 \times 25 \times 1 + 3 \times 5 \times 3 + 1 \times 5}{216}$$

$$= \frac{175}{216} = \$0.81$$

$\longrightarrow$ Avg. loss per game = \$0.19

(d) As part of a sales promotion, a breakfast cereal manufacturer randomly puts one of ten different toys in each packet of cereal. On average, how many packets would have to be purchased to get at least one of each different toy ?

For 1st toy $\rightarrow$ 1 packet is needed

For 2nd toy $\rightarrow \frac{10}{9}$ packets are needed on average

For 3rd toy $\rightarrow \frac{10}{8}$ packets are needed on average

etc . . . .

Total number of packets needed on average for all 10 toys

$$= 1 + \frac{10}{9} + \frac{10}{8} + \frac{10}{7} + \frac{10}{6} + \frac{10}{5} + \frac{10}{4} + \frac{10}{3} + \frac{10}{2} + \frac{10}{1}$$

$= 29.29 \longrightarrow \underline{30}$ because can buy only whole packets

Name ______________________ Student Number ______________

4. The probability density function for the hourly average wind speed v (in m/s) at an offshore "wind farm" is lognormal, meaning that x = ln(v) is normal

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right] \qquad \mu = 2.0;\ \sigma = 0.5$$

For low wind speed, power output P increases rapidly with wind speed. At moderate speeds, P is limited by the generator. At very high speeds, the wind turbine goes into survival mode and P=0. The function for P [kW] is:

v < 10	$P = Kv^3$ with K = 0.8 kW-s³/m³.
10 < v < 20	P = 800
v > 20	P = 0

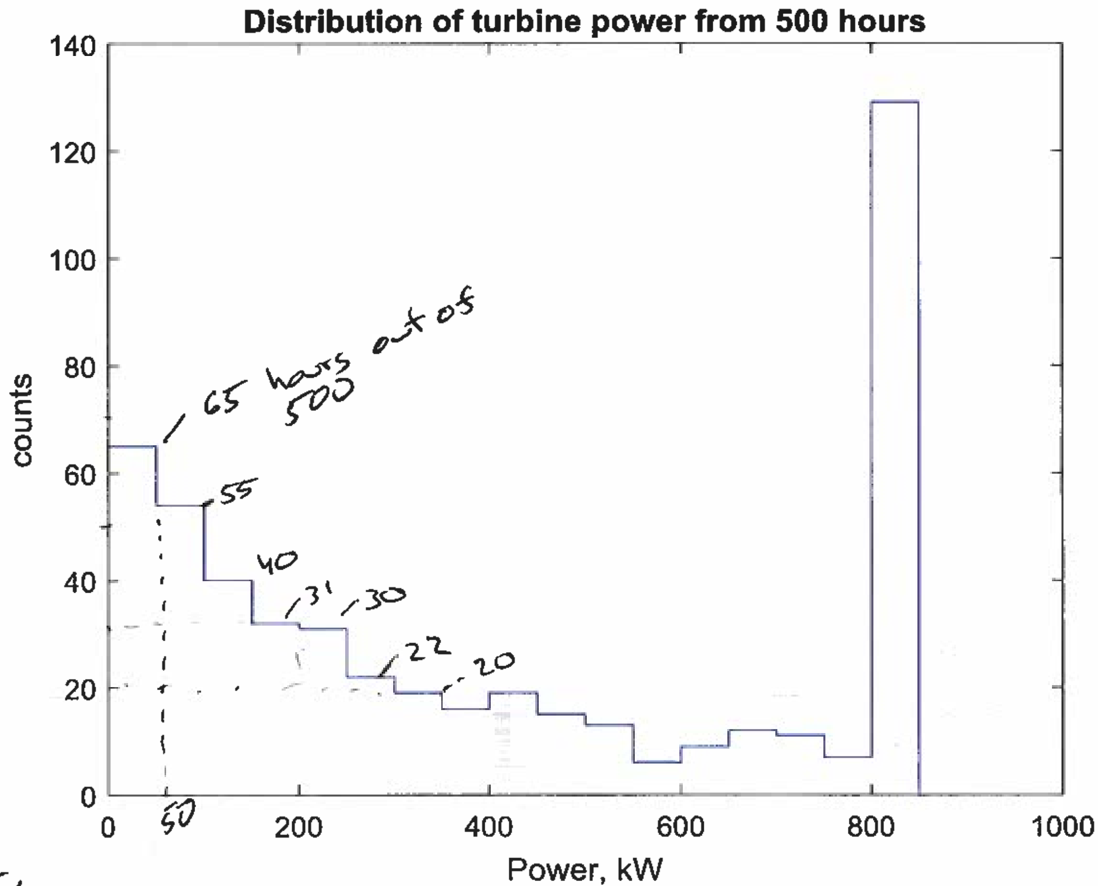
a) Write a Matlab script to produce a histogram of power output of the wind turbine. Try to use real Matlab functions and include comments on key lines to explain your logic.

```
clear all
medx = 2;
std = 0.5;
for i = 1:500
    z = randn;   % standard normal
    x = z* std + medx;
    V(i) = exp(x);        % velocity
    if V(i) < 10
        P(i) = 0.8 V(i)^3        % power, light winds
    else
        if V(i) < 20
            P(i) = 800 ;        % max power
        else
            P(i) = 0 ;          % survival
        end
    end
end
figure (1)
histogram(P, 20, 'Displaystyle', 'stairs', 'BinLimits', [0,1000]);
xlabel ('Power, kw')
ylable ('counts')
```

(this level of detail in Matlab formatting was not expected!)

b) The diagram below shows a histogram from this code for 500 simulated hours.

   i. Estimate the fraction of hours where P < 50 kW.
   ii. Determine the standard deviation of the result from question (i).
   iii. Why is there a large column at the right side of the histogram ?
   iv. Estimate the median power of the turbine.



i) 65/500

ii) 65 random counts, std = $\sqrt{N} \cong \underline{8}$

iii) all hours with $10 < V < 20$ m/s go to one bin

iv) compute running sum from left

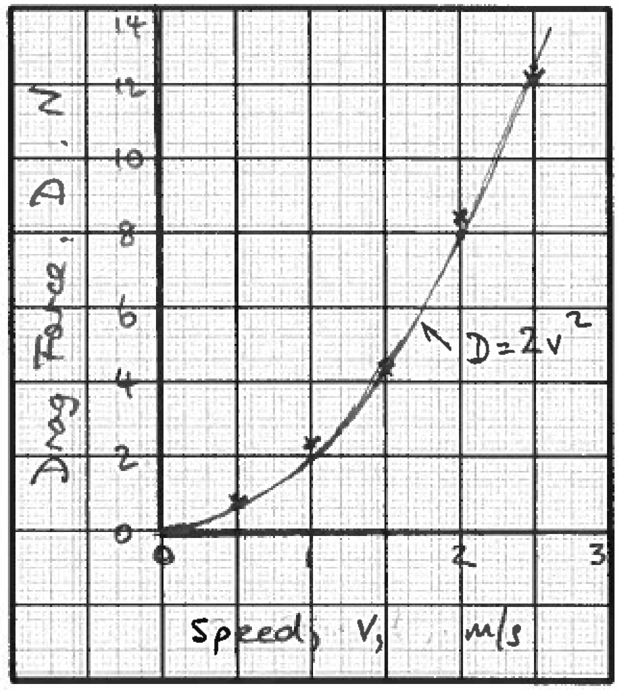
$\underline{P}$	Σ_i
50	65
100	120
150	160
200	191
250	221
300	243
350	263

250 hours arises for $300 < P < 350$ kW

so estimate median as ~ 325 kW

Name ______________________ Student Number ______________

5. A group of students went out on the Mech2 kayak field trip and measured the following data for the drag force on a kayak at various speeds:



Speed, V m/s	Drag Force, D N
0.5	0.7
1.0	2.3
1.5	4.3
2.0	8.4
2.5	12.1

a) Plot the data on the given graph paper.

b) It is known that the drag force D depends on speed V with a square-law relationship

$$D = K V^2$$

Use the least-squares method to determine the drag coefficient K. Explain your procedure in detail. *(Hint 1: the exponent 2 is known in advance and does not need to be statistically determined here. Hint 2: Do your calculations using at least three decimal places.)*

c) In an alternative approach, the drag force equation is rewritten as

$$d = k V$$

where $d = D^{0.5}$ and $k = K^{0.5}$. Use the least-squares method to determine k and hence a second estimate of K.

d) Likely your two estimates of K will be slightly different. Explain what significant features of the two computation methods causes the difference. Which method do you think may give a more realistic result ? Why ?

e) Plot one or other of your computed relationships, or some compromise intermediate relationship on your graph. Is the result reasonable ?

b) There is only one unknown, K. In matrix form, the data are:

$$\begin{bmatrix} V_1^2 \\ V_2^2 \\ V_3^2 \\ V_4^2 \\ V_5^2 \end{bmatrix} [K] = \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D4 \\ D_5 \end{bmatrix}$$

For least-squares solution premultiply by $\begin{bmatrix} V_1^2 \\ V_2^2 \\ V_3^2 \\ V_4^2 \\ V_5^2 \end{bmatrix}$

Name ______________________ Student Number ______________

$$\begin{bmatrix} V_1^2 & V_2^2 & V_3^2 & V_4^2 & V_5^2 \end{bmatrix} \begin{bmatrix} V_1^2 \\ V_2^2 \\ V_3^2 \\ V_4^2 \\ V_5^2 \end{bmatrix} \begin{bmatrix} K \end{bmatrix} = \begin{bmatrix} V_1^2 & V_2^2 & V_3^2 & V_4^2 & V_5^2 \end{bmatrix} \begin{bmatrix} D_1 \\ D_2 \\ D_3 \\ D_4 \\ D_5 \end{bmatrix}$$

$$\rightarrow \sum_{i=1}^{5} V_i^4 \, K = \sum_{i=1}^{5} V_i^2 D_i$$

$$\rightarrow K = \sum_{i=1}^{5} V_i^2 D_i \Big/ \sum_{i=1}^{5} V_i^4$$

V	D	V^4	V^2D
0.5	0.7	0.063	0.175
1.0	2.3	1.000	2.300
1.5	4.3	5.063	9.675
2.0	8.4	16.000	33.600
2.5	12.1	39.063	75.625
	sum	61.188	121.375

$$K = \frac{121.375}{61.188} = 1.984$$

$$\underline{D = 1.984\, V^2}$$

c) Repeat calculation using v and d instead of V^2 and D^2

$$\rightarrow k = \sum_{i=1}^{5} v_i d_i \Big/ \sum_{i=1}^{5} v_i^2$$

V	D	d	v^2	vd
0.5	0.7	0.837	0.250	0.418
1.0	2.3	1.517	1.000	1.517
1.5	4.3	2.074	2.250	3.110
2.0	8.4	2.898	4.000	5.797
2.5	12.1	3.479	6.250	8.696
		sum	13.750	19.538

$$k = \frac{19.538}{13.750} = 1.421$$

$$K = k^2 = 2.019$$

$$\underline{D = 2.019\, v^2}$$

Name ______________________________ Student Number ______________

(a blank page for your writing)

d) The least-squares method is based on the assumption that the measurement error is the same for all measurements. Thus, the largest measurements may be expected to dominate the calculation because the errors are smallest relative to the measurement size. This can be seen in the sums of quantities, where the early items are much smaller than the later items. In calculation (b), the errors are assumed to be the same for all D values, whereas in calculation (c), the error are assumed to be the same for all $d = \sqrt{D}$ values. The best-fit values of K are slightly different in each case.

I would guess that in reality the measurement error would increase with measurement size, so since calculation (c) gives less weighting to the larger values, it may be expected to give a more realistic result.

If it is expected that the errors are proportional to measurement size, then a logaritmic fit may be useful

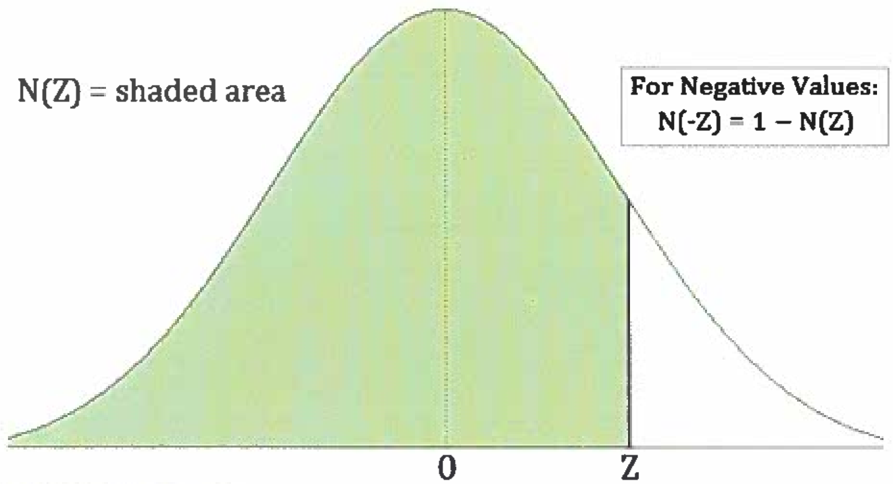
$$\ln D = \ln K + 2 \ln V$$

which gives $K = \exp\left(\left(\sum_{i=1}^{n} \ln D_i - 2 \sum_{i=1}^{n} \ln V_i\right) / n\right) = 2.19$

You could also try a logarithmic fit with two fitting parameters: K, and the power. In this case, I find K=2.32 and the power is 1.77 instead of 2. The model values for the 5 velocities are 0.68, 2.32, 4.77. 7.94 and 11.79.

Name ______________________________ Student Number ______________

Table of Values for N(Z)



Hundredths Digits

Tenths Digits	0.00	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0.0	0.5000	0.5040	0.5080	0.5120	0.5160	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0.5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5753
0.2	0.5793	0.5832	0.5871	0.5910	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0.6406	0.6443	0.6480	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.6700	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.6950	0.6985	0.7019	0.7054	0.7088	0.7123	0.7157	0.7190	0.7224
0.6	0.7257	0.7291	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.7580	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7852
0.8	0.7881	0.7910	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0.8315	0.8340	0.8365	0.8389
1.0	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0.8708	0.8729	0.8749	0.8770	0.8790	0.8810	0.8830
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.8980	0.8997	0.9015
1.3	0.9032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.9177
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.9370	0.9382	0.9394	0.9406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.9495	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0.9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1.9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.9750	0.9756	0.9761	0.9767
2.0	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.9808	0.9812	0.9817
2.1	0.9821	0.9826	0.9830	0.9834	0.9838	0.9842	0.9846	0.9850	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9875	0.9878	0.9881	0.9884	0.9887	0.9890
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0.9920	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2.5	0.9938	0.9940	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.9960	0.9961	0.9962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.9970	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0.9977	0.9978	0.9979	0.9979	0.9980	0.9981
2.9	0.9981	0.9982	0.9982	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986
3.0	0.9987	0.9987	0.9987	0.9988	0.9988	0.9989	0.9989	0.9989	0.9990	0.9990
3.1	0.9990	0.9991	0.9991	0.9991	0.9992	0.9992	0.9992	0.9992	0.9993	0.9993
3.2	0.9993	0.9993	0.9994	0.9994	0.9994	0.9994	0.9994	0.9995	0.9995	0.9995
3.3	0.9995	0.9995	0.9995	0.9996	0.9996	0.9996	0.9996	0.9996	0.9996	0.9997
3.4	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9997	0.9998