Midterm Examination Number One
Administered: Monday, February 8, 1999

In all relevant problems: WRITE DOWN THE FORMULA YOU USE, BEFORE YOU USE IT.
ALL PROBLEMS ARE WORTH 2 POINTS.

For Problems (1.1) to (1.7) Consider laminar flow down a circular pipe. 200 periodic readings of the maximum velocity (in the center of the pipe) and the viscosity were taken. We know that for laminar flow in a circular pipe that the average velocity [m/s] is:

$\overline{v} = \frac{v_{max}}{2}$, the mass flow rate [kg/s] is

$\dot{m} = \rho A\overline{v} = \rho \frac{\pi}{4} D^2 \overline{v}$, and the Reynold's number is

$N_{Re} = \frac{D\overline{v}\rho}{\mu}$. (Thus all three of these equations are functions of random variables, v_{max} and μ.)

The density, $\rho = 1000.0$ kg/m^3, and the diameter, $D = 0.01$ m. The following data is then tabulated for those 200 samples. (Not all of the raw data is shown below.)

run	vmax	mu	Nre	vmax^2	mu^2	Nre^2	vmax*Nre	m-flow-Nre	mu-Nre
	[m/s]	[kg/m/s]		[m/s]^2	[kg/m/s]^2		[m/s]	[kg/s]	[kg/m/s]
1	1.0411E+01	1.1663E-03	8.9266E+04	1.0838E+02	1.3601E-06	7.9685E+09	9.2933E+05	7.2989E+04	1.0411E+02
2	9.1656E+00	8.4819E-04	1.0806E+05	8.4008E+01	7.1942E-07	1.1677E+10	9.9045E+05	7.7789E+04	9.1656E+01
3	8.3684E+00	8.2060E-04	1.0198E+05	7.0030E+01	6.7339E-07	1.0400E+10	8.5340E+05	6.7026E+04	8.3684E+01
	…	…	…	…	…	…	…	…	…
198	1.1792E+01	1.1855E-03	9.9473E+04	1.3906E+02	1.4054E-06	9.8949E+09	1.1730E+06	9.2129E+04	1.1792E+02
199	9.1816E+00	1.0085E-03	9.1046E+04	8.4301E+01	1.0170E-06	8.2893E+09	8.3594E+05	6.5655E+04	9.1816E+01
200	9.5214E+00	8.9588E-04	1.0628E+05	9.0658E+01	8.0261E-07	1.1295E+10	1.0119E+06	7.9477E+04	9.5214E+01
sum	1.9934E+03	2.0227E-01	1.9994E+07	2.0097E+04	2.0734E-04	2.0518E+12	2.0162E+08	1.5835E+07	1.9934E+04
avg	9.9671E+00	1.0113E-03	9.9968E+04	1.0049E+02	1.0367E-06	1.0259E+10	1.0081E+06	7.9175E+04	9.9671E+01

(1.1) Find the variance of the average velocity, $\sigma^2_{\overline{v}}$.

(1.2) Find the standard deviation of the viscosity, σ_{μ}.

(1.3) Find the mean of the mass-flow, $\mu_{\dot{m}}$.

(1.4) Find the variance of the Reynold's number, $\sigma^2_{N_{Re}}$.

(1.5) Find the covariance of the viscosity and Reynold's number, $\sigma_{\mu N_{Re}}$.

(1.6) Find the correlation coefficient, $\rho_{\mu N_{Re}}$.

(1.7) Explain the sign of $\rho_{\mu N_{Re}}$.

Solution:

(1.1) Find the variance of the average velocity, $\sigma^2_{\overline{v}}$.

$$\sigma_{\bar{v}}^2 = E\left(\bar{v}^2\right) - E\left(\bar{v}\right)^2 = E\left(\left(\frac{v_{max}}{2}\right)^2\right) - E\left(\frac{v_{max}}{2}\right)^2 = \frac{1}{4}E\left({v_{max}}^2\right) - \frac{1}{4}E\left(v_{max}\right)^2$$

$$\sigma_{\bar{v}}^2 = \frac{1}{4}\cdot 100.49 - \frac{1}{4}9.9671^2 = 0.287 \text{ m}^2/\text{s}^2$$

(1.2) Find the standard deviation of the viscosity, σ_μ.

$$\sigma_\mu = \sqrt{\sigma_\mu^2} = \sqrt{E\left(\mu^2\right) - E\left(\mu\right)^2} = \sqrt{1.0367\cdot 10^{-6} - 1.0113\cdot {10^{-3}}^2}$$

$$\sigma_\mu = 1.18\cdot 10^{-4} \text{ kg/m/s}$$

(1.3) Find the mean of the mass-flow, $\mu_{\dot{m}}$.

$$\mu_{\dot{m}} = E\left(\dot{m}\right) = E\left(\rho\frac{\pi}{4}D^2\bar{v}\right) = E\left(\rho\frac{\pi}{4}D^2\frac{v_{max}}{2}\right) = \rho\frac{\pi}{8}D^2E\left(v_{max}\right)$$

$$\mu_{\dot{m}} = 1000.0\frac{\pi}{8}0.01^2\cdot 9.9671 = 0.391 \text{ kg/s}$$

(1.4) Find the variance of the Reynold's number, $\sigma_{N_{Re}}^2$.

$$\sigma_{N_{Re}}^2 = E\left({N_{Re}}^2\right) - E\left(N_{Re}\right)^2 = 1.0259\cdot 10^{10} - 9.9968\cdot {10^4}^2 = 2.65\cdot 10^8$$

(1.5) Find the covariance of the viscosity and Reynold's number, $\sigma_{\mu N_{Re}}$.

$$\sigma_{\mu N_{Re}} = E\left(\mu N_{Re}\right) - E\left(\mu\right)E\left(N_{Re}\right) = 99.671 - \left(1.0113\cdot 10^{-3}\right)\left(9.9968\cdot 10^4\right)$$

$$\sigma_{\mu N_{Re}} = -1.4266 \text{ kg/m/s}$$

(1.6) Find the correlation coefficient, $\rho_{\mu N_{Re}}$.

$$\rho_{\mu N_{Re}} = \frac{\sigma_{\mu N_{Re}}}{\sqrt{\sigma_\mu^2}\sqrt{\sigma_{N_{Re}}^2}} = \frac{-1.4266}{1.18\cdot 10^{-4}\sqrt{2.65\cdot 10^8}} = -0.74$$

(1.7) Explain the sign of $\rho_{\mu N_{Re}}$.

The negative sign of $\rho_{\mu N_{Re}}$ indicates that as the viscosity increases, the Reynold's decreases and that as the viscosity decreases, the Reynold's increases. This is what we expect from the definition of the Reynold's number.

For problems (2.1) to (2.4) consider the Joint PDF $f(x,y)$ given by:

y	x→	1	2	3	4
2		1/12	3/24	1/24	1/6
4		1/24	1/12	1/24	1/24

6	1/6	1/12	1/12	1/24

(2.1) Is this PDF discrete or continuous?
(2.2) Show this PDF is a valid PDF.
(2.3) Find $P(X \geq 4, Y = 4)$
(2.4) Find the marginal distribution, h(y)
(2.5) Find $P(X \geq 4 \mid Y = 4)$

Solution:

(2.1) Is this PDF discrete or continuous?
Discrete.

(2.2) Show this PDF is a valid PDF.
The sum of the probabilities is one.
No probability is negative.
The PDF is valid.

(2.3) Find $P(X \geq 4, Y = 4)$

$$P(X \geq 4, Y = 4) = P(X = 4, Y = 4)$$
$$P(X \geq 4, Y = 4) = 1/24$$

(2.4) Find the marginal distribution, h(y)

$$h(Y = 2) = 10/24$$
$$h(Y = 4) = 5/24$$
$$h(Y = 6) = 9/24$$

(2.5) Find $P(X \geq 4 \mid Y = 4)$

$$P(X \geq 4 \mid Y = 4) = \frac{P(X \geq 4, Y = 4)}{h(Y = 4)} = \frac{1/24}{5/24} = \frac{1}{5}$$

For questions (3.1) to (3.6):
In sampling a population for the presence of a disease, the population is of two types: **I**nfected and **U**ninfected. The results of the test are of two types: **P**ositive and **N**egative. In rare disease detection, a high probability for detecting a disease can still lead to more false positives than true positives. Consider a case where a disease infects 1 out of every 100,000 individuals. The probability for a positive test result given that the subject is infected is 0.99. The probability for a negative test result given that the subject is uninfected is 0.999.

(3.1) For testing a single person, define the complete sample space.

(3.2) What is the probability of a false negative test result (a negative test result given that the subject is infected)?

(3.3) What is the probability of being uninfected AND having a negative test result?

(3.4) What is the probability of testing positive?

(3.5) Determine rigorously whether testing positive and having the disease are independent.

(3.6) Determine the percentage of people who test positive who are really uninfected.

(3.7) In a population of 250 million, with the infection rate given, how many people would you expect to be (a) Infected-test Positive, (b) Infected-test Negative, (c) Uninfected-test Positive, (d) Uninfected-test negative.

Solution:

We are told:

$$P(I) = 10^{-5}$$

$$P(N|U) = \frac{P(N \cap U)}{P(U)} = 0.999$$

$$P(P|I) = \frac{P(P \cap I)}{P(I)} = 0.99$$

(3.1) For testing a single person, define the complete sample space.

The sample space is $S = \{IP, IN, UP, UN\}$ where I = Infected, U=Uninfected, P=positive test result, N=negative test result.

(3.2) What is the probability of a false negative test result (a negative test result given that the subject is infected)?

We want: $P(N|I) = \frac{P(N \cap I)}{P(I)}$ so we need the two factors on the right hand side.

We have been given the denominator. In order to find the numerator, we must use the other given:

$$P(P|I) = \frac{P(P \cap I)}{P(I)} = 0.99$$

which rearranges for the intersection of P AND I

$$P(P \cap I) = P(I) \cdot P(P|I) = (10^{-5})(0.99) = 0.99 \cdot 10^{-5}$$

We must realize that the probability of I is the union of IP and IN groups.
So using the definition of the Union, we have:

$$P(I) = P[(I \cap P) \cup (I \cap N)] = P(I \cap P) + P(I \cap N) - P[(I \cap P) \cap (I \cap N)]$$

The result cannot be both positive and negative:

$$P[(I \cap P) \cap (I \cap N)] = 0$$

So,

$$P(I \cap N) = P(I) - P(I \cap P) = (10^{-5}) - 0.99 \cdot 10^{-5} = 10^{-7}$$

Then we can plug into our original equation:

$$P(N|I) = \frac{P(N \cap I)}{P(I)} = \frac{10^{-7}}{10^{-5}} = 0.01$$

OR, an alternative solution, relies on us recognizing:

$P(N|I) + P(P|I) = 1$ because every test comes out positive or negative.

$$P(N|I) = 1 - P(P|I) = 1 - 0.99 = 0.01$$

(3.3) What is the probability of being uninfected AND having a negative test result?

We want $P(N \cap U)$

we can obtain this from either:

(a) the UNION RULE:

$$P(N) = P[(N \cap I) \cup (N \cap U)]$$
$$P(N) = P(N \cap I) + P(N \cap U) - P[(N \cap I) \cap (N \cap U)]$$
$$P[(N \cap I) \cap (N \cap U)] = 0$$
$$P(N) = P(N \cap I) + P(N \cap U)$$
$$P(N \cap U) = P(N) - P(N \cap I)$$

but we don't know $P(N \cap I)$ and we don't know $P(N)$

or (b) the conditional probability rule:

$$P(U|N) = \frac{P(N \cap U)}{P(N)}$$
$$P(N \cap U) = P(N) \cdot P(U|N)$$

but we don't know $P(U|N)$ and we don't know $P(N)$

or (c) the conditional probability rule:

$$P(N|U) = \frac{P(N \cap U)}{P(U)} = 0.999$$
$$P(N \cap U) = P(U) \cdot P(N|U) = P(U) \cdot 0.999$$

I like choice (c) because we are given $P(N|U) = 0.999$ and we know

$P(U) = 1 - P(I) = 1 - 10^{-5} = 0.99999$ so

$$P(N \cap U) = P(U) \cdot P(N|U) = (0.99999)(0.999) = 0.99899001$$

(3.4) What is the probability of testing positive?

We want $P(P)$

We can find $P(P)$ either by:

(a) the fact that the sum of the probabilities must be one

$P(P) + P(N) = 1$ but we don't know $P(N)$

$$P(P) = 1 - P(N)$$

(b) the conditional probability distribution:

$$P(I|P) = \frac{P(P \cap I)}{P(P)}$$ but we don't know $P(I|P)$

(c) the conditional probability distribution:

$$P(U|P) = \frac{P(P \cap U)}{P(P)}$$ but we don't know $P(U|P)$

(d) the sum of the probabilities must be one and a different conditional probability:

$$P(P) = 1 - P(N)$$

$$P(U|N) = \frac{P(N \cap U)}{P(N)}$$

$$P(P) = 1 - P(N) = 1 - \frac{P(N \cap U)}{P(U|N)}$$ but we don't know $P(U|N)$

(e) the sum of the probabilities must be one and a different conditional probability:

$$P(P) = 1 - P(N)$$

$$P(I|N) = \frac{P(N \cap I)}{P(N)}$$

$$P(P) = 1 - P(N) = 1 - \frac{P(N \cap I)}{P(I|N)}$$ but we don't know $P(I|N)$

(f) the Union rule:

$$P(P) = P[(P \cap I) \cup (P \cap U)]$$

$$P(P) = P(P \cap I) + P(P \cap U) - P[(P \cap I) \cap (P \cap U)]$$

$$P[(P \cap I) \cap (P \cap U)] = 0$$

$$P(P) = P(P \cap I) + P(P \cap U)$$

combine with conditional probabilities that we do know:

$$P(P) = P(I) * P(P | I) + P(U) * P(P | U)$$

I like choice (f):

$$P(P) = 10^{-5} \cdot 0.99 + 0.99999 \cdot P(P | U)$$

we can get the last factor by considering (as we did in part (1.2))

$$P(P|U) + P(N|U) = 1$$ because all tests are either positive or negative.

$$P(P|U) = 1 - P(N|U) = 1 - 0.999 = 0.001$$

so

$$P(P) = 10^{-5} \cdot 0.99 + 0.99999 \cdot 0.001 = 0.00100989$$

(3.5) Determine rigorously whether testing positive and having the disease are independent.

If $P(P)$ and $P(I)$ are independent:

$$P(P \cap I) = P(P) \cdot P(I)$$

$$0.99 \cdot 10^{-5} = 0.00100989 \cdot 10^{-5}$$

NOT INDEPENDENT.

(3.6) Determine the percentage of people who test positive but who are really uninfected.

We want: $\frac{P(P \cap U)}{P(P)}$

$$\frac{P(P \cap U)}{P(P)} = \frac{0.99999 \cdot 10^{-3}}{0.00100989} = 0.990196952 = 99\%$$

<u>Despite the high accuracy of the test 99% of those people who test positive are actually uninfected.</u>

NOT INCLUDED IN EXAM.
(3.7) In a population of 250 million, with the infection rate given, how many people would you expect to be (a) Infected-test Positive, (b) Infected-test Negative, (c) Uninfected-test Positive, (d) Uninfected-test negative.

These are four intersections:

From part (3.5) we know:

$$P(P \cap I) = 0.99 \cdot 10^{-5}$$

$$P(P \cap U) = 0.99999 \cdot 0.001 = 0.99999 \cdot 10^{-3}$$

From part (3.2) we know

$$P(N \cap I) = P(I) - P(P \cap I) = \left(10^{-5}\right) - 0.99 \cdot 10^{-5} = 10^{-7}$$

From part (3.3) we know

$$P(N \cap U) = P(U) \cdot P(N|U) = (0.99999)(0.999) = 0.99899001$$

These should sum to 1.0 and they do.

Out of 250 million people, the number who are infected and test positive are: 2475.
Out of 250 million people, the number who are infected and test negative are: 25.
Out of 250 million people, the number who are uninfected and test positive are: 249,997.5
Out of 250 million people, the number who are uninfected and test negative are: 249.7475 million