IE241 HW1 Solution, Total = 35 pts

1.11 The following results on summations will help us in calculating the sample variance s^2. For any constant c, (5pts)

a $\sum_{i=1}^{n} c = nc$.

b $\sum_{i=1}^{n} cy_i = c\sum_{i=1}^{n} y_i$.

c $\sum_{i=1}^{n} (x_i + y_i) = \sum_{i=1}^{n} x_i + \sum_{i=1}^{n} y_i$.

Use (a), (b), and (c) to show that

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2\right].$$

Pr) $s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \sum_{i=1}^{n} 2y_i\bar{y} + \sum_{i=1}^{n}\bar{y}^2\right]$ by (c).

$= \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - 2\bar{y}\sum_{i=1}^{n} y_i + n\bar{y}^2\right]$ by (a), (b)

$= \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - 2\bar{y}\cdot(n\bar{y}) + n\bar{y}^2\right] = \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - n\bar{y}^2\right]$

$= \frac{1}{n-1}\left[\sum_{i=1}^{n} y_i^2 - \frac{1}{n}\left(\sum_{i=1}^{n} y_i\right)^2\right].$

1.22 Prove that the sum of the deviations of a set of measurements about their mean is equal to zero; that is,

$$\sum_{i=1}^{n}(y_i - \bar{y}) = 0.$$

$\sum_{i=1}^{n} y_i - n\bar{y} = 0$, AS you know $\bar{y} = \frac{1}{n}\sum y_i$

$\therefore\ n\bar{y} - n\bar{y} = 0$

1.27 A set of 340 examination scores exhibiting a bell-shaped relative frequency distribution has a mean of $\bar{y} = 72$ and a standard deviation of $s = 8$. Approximately how many of the scores would you expect to fall in the interval from 64 to 80? The interval from 56 to 88?

If data are bell-shape, we can use empirical rule (see. 2/27 note p.g 10)

$(64, 88) = (\bar{y} - 1sd, \bar{y} + 1sd) = (72-8, 72+8)$ contains 68% $\Rightarrow 340 \times 0.68 \simeq 231$

$(56, 88) = (\bar{y} - 2sd, \bar{y} + 2sd) = (72-16, 72+16)$ contains 95% $\Rightarrow 340 \times 0.95 \simeq 323$

*1.32 Let $k \geq 1$. Show that, for any set of n measurements, the fraction included in the interval $\bar{y} - ks$ to $\bar{y} + ks$ is at least $(1 - 1/k^2)$. [*Hint:*

$$s^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i - \bar{y})^2\right].$$

In this expression, replace all deviations for which $|y_i - \bar{y}| \geq ks$ with ks. Simplify.] This result is known as *Tchebysheff's theorem*.[3]

(10pts)

pr) Let $A = \{i : |y_i - \bar{y}| \geq ks\}$, n_A = # of measurements that outside the interval.
$A^c = \{i : |y_i - \bar{y}| < ks\}$, n_{A^c} = // inside the interval.
Fraction included in the interval $= \frac{n - n_A}{n} = \frac{n_{A^c}}{n}$

$$\sum_{i=1}^{n}(y_i - \bar{y})^2 = \sum_{A}(y_i - \bar{y})^2 + \sum_{A^c}(y_i - \bar{y})^2$$
$$\geq n_A k^2 s^2 + \sum_{A^c}(y_i - \bar{y})^2 \quad \text{using } (y_i - \bar{y})^2 \geq k^2 s^2$$
$$\geq n_A k^2 s^2 \quad \text{using } \sum_{A^c}(y_i - \bar{y})^2 \geq 0$$

$\therefore \; (n-1)s^2 \geq n_A \cdot k^2 \cdot s^2$.

$ns^2 \geq (n-1)s^2 \geq n_A \cdot k^2 \cdot s^2 \iff n \geq n_A \cdot k^2 \iff \frac{1}{k^2} \geq \frac{n_A}{n}$ using this we can get

$$1 - \frac{1}{k^2} \leq \frac{n - n_A}{n} = \frac{n_{A^c}}{n}$$

1.37 Studies indicate that drinking water supplied by some old lead-lined city piping systems may contain harmful levels of lead. Based on data presented by Karalekas and colleagues,[4] it appears that the distribution of lead content readings for individual water specimens has mean .033 mg/L and standard deviation .10 mg/L. Explain why it is obvious that the lead content readings are *not* normally distributed.

normal distribution → symmetric & bell shape
If data are normal, almost 68% of data included in 0.033 ± 0.1
But data can only have non-negative values. Hence Not normally distributed.

2.18 Suppose two balanced coins are tossed and the upper faces are observed.

a List the sample points for this experiment.

b Assign a reasonable probability to each sample point. (Are the sample points equally likely?)

c Let A denote the event that *exactly* one head is observed and B the event that *at least* one head is observed. List the sample points in both A and B.

d From your answer to part (c), find $P(A)$, $P(B)$, $P(A \cap B)$, $P(A \cup B)$, and $P(\bar{A} \cup B)$.

a. $S = \{HH, HT, TH, TT\}$
b. two balance coins → all events have same probability : $\frac{1}{4}$
c. $A = \{HT, TH\}$
$B = \{HT, TH, HH\}$

d. $P(A) = 0.5$, $P(B) = 0.75$, $P(A \cap B) = 0.5$, $P(A \cup B) = 0.75$, $P(A^c \cup B) = 1$

→ This problem well Known as 'Monti Hall problem'

2.20 The following game was played on a popular television show. The host showed a contestant three large curtains. Behind one of the curtains was a nice prize (maybe a new car) and behind the other two curtains were worthless prizes (duds). The contestant was asked to choose one curtain. If the curtains are identified by their prizes, they could be labeled G, D_1, and D_2 (Good Prize, Dud1, and Dud2). Thus, the sample space for the contestants choice is $S = \{G, D_1, D_2\}$.[1] (10pts)

a If the contestant has no idea which curtains hide the various prizes and selects a curtain at random, assign reasonable probabilities to the simple events and calculate the probability that the contestant selects the curtain hiding the nice prize.

b Before showing the contestant what was behind the curtain initially chosen, the game show host would open one of the curtains and show the contestant one of the duds (he could always do this because he knew the curtain hiding the good prize). He then offered the contestant the option of changing from the curtain initially selected to the other remaining unopened curtain. Which strategy maximizes the contestant's probability of winning the good prize: stay with the initial choice or switch to the other curtain? In answering the following sequence of questions, you will discover that, perhaps surprisingly, this question can be answered by considering only the sample space above and using the probabilities that you assigned to answer part (a).

i If the contestant choses to stay with her initial choice, she wins the good prize if and only if she initially chose curtain G. If she stays with her initial choice, what is the probability that she wins the good prize?

ii If the host shows her one of the duds and she switches to the other unopened curtain, what will be the result if she had initially selected G?

iii Answer the question in part (ii) if she had initially selected one of the duds.

iv If the contestant switches from her initial choice (as the result of being shown one of the duds), what is the probability that the contestant wins the good prize?

v Which strategy maximizes the contestant's probability of winning the good prize: stay with the initial choice or switch to the other curtain?

(a) The contestant has no idea → has same probability

$\therefore P(G) = P(D_1) = P(D_2) = \frac{1}{3}$

(b). I The contestant choose to stay her initial choice.

Probability of selecting G at her initial Choice = Probability of getting the good prize

$\therefore \frac{1}{3}$

II. Initially select G → switch curtain → Get D_1 or D_2

III. 〃 D_1 or D_2 → 〃 → Get G

IV. Since she has decided to change her first choice, she gets a good prize if she chooses D_1 or D_2 in her first choice $\Rightarrow P = \frac{2}{3}$

V. By I, IV, the best strategy is to switch.

2.28 Four equally qualified people apply for two identical positions in a company. One and only one applicant is a member of a minority group. The positions are filled by choosing two of the applicants at random.

a List the possible outcomes for this experiment.

b Assign reasonable probabilities to the sample points.

c Find the probability that the applicant from the minority group is selected for a position.

Four People = $\{A_1, A_2, A_3, \tilde{A}_4\}$

↳ minority group

a. $S = \{A_1A_2, A_1A_3, A_1\tilde{A}_4, A_2A_3, A_2\tilde{A}_4, A_3\tilde{A}_4\}$

b. All cases have same Probability $= \frac{1}{6}$

c. $P(A_1\tilde{A}_4) + P(A_2\tilde{A}_4) + P(A_3\tilde{A}_4) = \frac{1}{2}$

2.48 If we wish to expand $(x + y)^8$, what is the coefficient of x^5y^3? What is the coefficient of x^3y^5?

As you know $(x+y)^n = \sum_{i=0}^{n} \binom{n}{i} x^{n-i} \cdot y^i$: binomial theorem

Hence Coef of x^5y^3 = Coef of x^3y^5 = $\binom{8}{3}$ or $\binom{8}{5}$

$\therefore \binom{8}{3} = \binom{8}{5} = \frac{8!}{3!\,5!} = 56$

2.64 A balanced die is tossed six times, and the number on the uppermost face is recorded each time. What is the probability that the numbers recorded are 1, 2, 3, 4, 5, and 6 in any order? (5pts)

Total possible cases (sample space) $= 6^6$

Number recorded are 1, 2, 3, 4, 5, 6 in any order $\Rightarrow 6!$

$\rightarrow \frac{6!}{6^6} = \frac{5}{324}$

2.68 Show that, for any integer $n \geq 1$, (5pts)

a $\binom{n}{n} = 1$. Interpret this result.

b $\binom{n}{0} = 1$. Interpret this result.

c $\binom{n}{r} = \binom{n}{n-r}$. Interpret this result.

d $\sum_{i=0}^{n} \binom{n}{i} = 2^n$. [*Hint:* Consider the binomial expansion of $(x + y)^n$ with $x = y = 1$.]

a. $\binom{n}{n} = \frac{n!}{n!\,\underbrace{(n-n)!}_{=0!=1}} = 1$. only one way to choose all

b. $\binom{n}{0} = \frac{n!}{0!\,(n-0)!} = 1$. only one way to chose nothing

c. $\binom{n}{r} = \frac{n!}{r!\,(n-r)!} = \binom{n}{n-r} = \frac{n!}{(n-r)!\,(n-(n-r))!}$. Choose r out of n = Choose (n-r) out of n

d. by 2.48, $2^n = \sum_{i=0}^{n} \binom{n}{i} \cdot 1^{n-i} \cdot 1^i = \sum_{i=0}^{n} \binom{n}{i}$

2.72 For a certain population of employees, the percentage passing or failing a job competency exam, listed according to sex, were as shown in the accompanying table. That is, of all the people taking the exam, 24% were in the male-pass category, 16% were in the male-fail category, and so forth. An employee is to be selected randomly from this population. Let A be the event that the employee scores a passing grade on the exam and let M be the event that a male is selected.

	Sex		
Outcome	Male (M)	Female (F)	*Total*
Pass (A)	24	36	60
Fail ($\overline{A}$)	16	24	40
Total	40	60	100

a Are the events A and M independent?

b Are the events $\overline{A}$ and F independent?

a. $P(A) = \frac{60}{100}$, $P(M) = \frac{40}{100}$ / $P(A)P(M) = 0.24 = P(A \cap M) = 0.24$

$\Rightarrow$ Independent

b. $P(\overline{A}) = \frac{40}{100}$, $P(F) = \frac{60}{100}$ / $P(\overline{A})P(F) = 0.24 = P(\overline{A} \cap F) = 0.24$

$\Rightarrow$ Independent