MULTIPLE CHOICE

- A probability distribution of the claim sizes (in hundreds of dollars) for an automobile insurance policy is given in the table below:

Claim size	20	30	40	50	60	70	80
Probabililty	0.25	0.10	0.05	0.20	0.10	0.10	0.20

What percentage of the claims are within one standard deviation of the mean claim size?

We calculate: $\mu = 20(0.05) + 30(0.10) + \cdots + 70(0.10) + 80(0.2) = 49$

$$\sigma^2 = 20^2(0.05) + 30^2(0.10) + \cdots + 70^2(0.10) + 80^2(0.2) - 49^2$$
$$= 2900 - 2401$$
$$= 499$$

Thus, $\sigma = \sqrt{499} = 22.34$ and so $(\mu - \sigma, \mu + \sigma) = (26.66, 71.34)$

Desired probability is $0.10 + 0.05 + 0.20 + 0.10 + 0.10 = 0.55$ or 55%

- The probability that a visit to a primary care physician's (PCP) office results in neither lab work nor referral to a specialist is 0.35. Of those coming to a PCP's office, 30% are referred to specialists and 40% require lab work. What is the probability that a visit to a PCP's office results in both lab work and referral to a specialist?

Let $A \equiv$ referral to lab work
$B \equiv$ referral to specialist

Given: $P(\overline{A \cup B}) = 0.35$
$P(A) = 0.4$
$P(B) = 0.3$

By the Addition Rule,
$$P(A \cup B) = P(A) + P(B) - P(A \cap B)$$
$$P(A \cap B) = P(A) + P(B) - P(A \cup B)$$
$$P(A \cap B) = P(A) + P(B) - 1 + P(\overline{A \cup B}) \text{ by Subtraction Rule}$$
$$P(A \cap B) = 0.4 + 0.3 - 1 + 0.35 = 0.05$$

- You are given $P(A \cup B) = 0.7$ and $P(A \cup \overline{B}) = 0.9$. What is $P(A)$?

Note that $(A \cup B) \cup (A \cup \overline{B}) = S$ since $B \cup \overline{B} = S$.

By the Addition Rule,
$$P(S) = P(A \cup B) + P(A \cup \overline{B}) - P(\{A \cup B\} \cap \{A \cup \overline{B}\})$$
$$1 = 0.7 + 0.9 - P(A) \text{ since } A \text{ is the only common piece}$$
$$P(A) = 1.6 - 1$$
$$P(A) = 0.6$$

- Suppose a fair die is rolled once. Which of the following pairs of events are mutually exclusive?

By definition of "mutually exclusive", the two events which share no common outcomes are: $A \equiv$ the numbers greater than 4; $B \equiv$ the numbers less than 3

- For a given data set, suppose that $\bar{x} = 13$ and $s^2 = 9$. According to Tchebysheff's Theorem, a lower bound for the proportion of data points in the interval $(9, 17)$ is

We need to find k such that $(\bar{x}-ks, \bar{x}+ks) = (9, 17)$

$$(13-3k, 13+3k) = (9, 17)$$

and so $13-3k=9$ or $13+3k=17$

$3k=4$ $\quad$ $3k=4$

$k=4/3$ $\quad$ $k=4/3$

By Tchebysheff's Theorem, a lower bound is given by $1-\frac{1}{k^2} = 1-\left(\frac{3}{4}\right)^2$

$$= 1-\frac{9}{16}$$

$$= \frac{7}{16}$$

$$\simeq 0.44$$

- A discrete random variable X can only take values in $\{1, 2, 3, 4\}$ with probability function $p(x) = k(2x-1)$ for some (unknown) value k. The probability that $X = 3$ is given by

Since $\sum_{\text{all } x} p(x) = 1$, we must have $p(1)+p(2)+p(3)+p(4)=1$

$$k+3k+5k+7k=1$$

$$16k=1$$

$$k=1/16$$

Thus, $P(X=3) = p(3) = \frac{2(3)-1}{16} = 5/16$

- The police in a small community know that 20% of the homeowners leave their doors unlocked. Crime records show that 10% of the homes whose doors are locked are burglarized while 75% of those with unlocked doors are burglarized. What is the probability that a burglarized home was found to have its doors unlocked?

Let $B \equiv$ burglarized home

$U \equiv$ unlocked home

Given: $P(U)=0.2$

$P(B|\bar{U})=0.1$

$P(B|U)=0.75$

Using Bayes' Theorem, we want $P(U|B) = \frac{P(U)P(B|U)}{P(U)P(B|U)+P(\bar{U})P(B|\bar{U})}$

$$= \frac{(0.2)(0.75)}{(0.2)(0.75)+(0.8)(0.1)}$$

$$= \frac{0.15}{0.15+0.08}$$

$$= \frac{0.15}{0.23} \simeq 0.65$$

- Weekly CPU time (measured in hours) used by a local accounting firm is a random variable having mean 55 and variance 196. The CPU time costs the firm \$200 per hour plus a \$50 overhead charge. What is the standard deviation of the weekly cost for CPU time?

Let X represent the weekly CPU time (measured in hours).

Cost random variable is $C = 200X+50$, which is linear in X.

Thus, $\sigma_C = 200\sigma_X = 200\sqrt{196} = 2800$

1. A recent financial study examined the fluctuation in the monthly interest rate of two particular mutual funds (i.e. Fund A and Fund B) offered by the Bank of Montreal. The following table gives the end-of-month interest rates of both of these mutual funds for 7 randomly selected months in 2009:

Fund A (x)	7.31	10.50	5.86	8.82	16.11	7.87	7.22
Fund B (y)	3.17	3.46	3.05	3.36	3.55	3.25	3.12

(4 marks) (a) Calculate the median and IQR of the Fund A data.

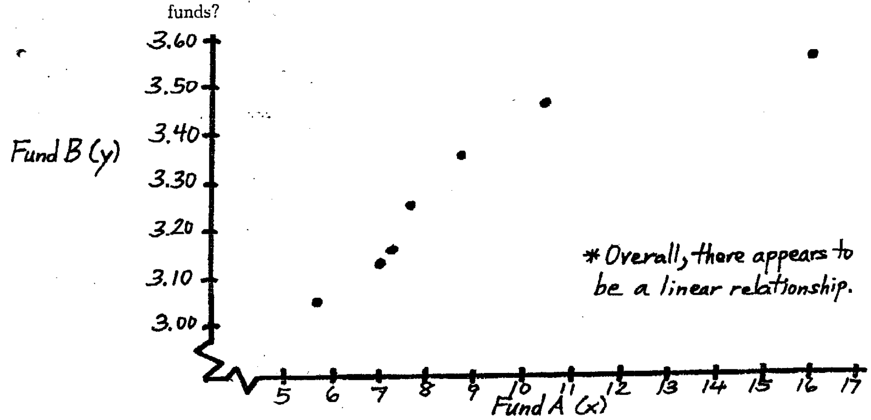
First, order the observations:

5.86 7.22 7.31 7.87 8.82 10.50 16.11

$\tilde{x} = \text{median} = 7.87$

$$IQR = Q_3 - Q_1 = \frac{8.82+10.50}{2} - \frac{7.22+7.31}{2} = 9.66 - 7.265 = 2.395$$

(4 marks) (b) Construct a scatter plot of the above data. Does there appear to be a linear relationship between the interest rates of these two funds?



* Overall, there appears to be a linear relationship.

(4 marks) (c) Compute the correlation coefficient and interpret it in the context of the question.

We need to compute: $\sum_{i=1}^{7} x_i = 7.31 + \dots + 7.22 = 63.69$

$$\sum_{i=1}^{7} y_i = 3.17 + \dots + 3.12 = 22.96$$

$$\sum_{i=1}^{7} x_i y_i = (7.31)(3.17) + \dots + (7.22)(3.12) = 212.3053$$

$$\sum_{i=1}^{7} x_i^2 = (7.31)^2 + \dots + (7.22)^2 = 649.416$$

$$\sum_{i=1}^{7} y_i^2 = (3.17)^2 + \dots + (3.12)^2 = 75.512$$

Then, $S_x = \sqrt{\frac{1}{6}\left[649.416 - \frac{(63.69)^2}{7}\right]} = 3.414$

$$S_y = \sqrt{\frac{1}{6}\left[75.512 - \frac{(22.96)^2}{7}\right]} = 0.184$$

Thus, $Cor(x,y) = \dfrac{\frac{1}{6}\left[212.3053 - \frac{(63.69)(22.96)}{7}\right]}{(3.414)(0.184)} = 0.903$

* The correlation is very close to $+1$, indicating that there is strong evidence of a **POSITIVE** linear association between the 2 interest rates.

2. The annual rates of return on 12 randomly sampled foreign stocks are as follows:

8.9 3.1 4.7 2.8 9.0 7.1 5.7 5.3 5.2 6.6 5.2 5.7

[3 marks] (a) Construct an appropriate stem and leaf plot for the above data.

First, order the observations:

2.8	3.1	4.7	5.2	5.2	5.3
5.7	5.7	6.6	7.1	8.9	9.0

Stem and Leaf Plot:

Stem = ones digit
Leaf = tenths digit

2	8
3	1
4	7
5	22377
6	6
7	1
8	9
9	0

[3 marks] (b) Calculate (to 3 decimal places of accuracy) the mean and standard deviation of this data.

$$\bar{x} = \frac{1}{12}(8.9 + 3.1 + \cdots + 5.7) = \frac{69.3}{12} = 5.775$$

$$S = \sqrt{\frac{1}{12-1}\left[(8.9^2 + 3.1^2 + \cdots + 5.7^2) - \frac{(69.3)^2}{12}\right]}$$

$$= \sqrt{\frac{1}{11}(440.87 - 400.2075)}$$

$$= 1.923$$

[4 marks] (c) Does this data conform to the Empirical Rule? Justify your answer.

Look at the intervals: $(\bar{x}-s, \bar{x}+s) = (5.775 - 1.923, 5.775 + 1.923)$

$= (3.852, 7.698)$

$8/12 = 66.7\%$ of the data lie in this interval

$(\bar{x}-2s, \bar{x}+2s) = (5.775 - 3.846, 5.775 + 3.846)$

$= (1.929, 9.621)$

$12/12 = 100\%$ of the data lie in this interval

*Empirical Rule says approx. 68% and 95% should lie in these intervals $\Longrightarrow$ Based on our findings, data does not quite conform.

1. A recent financial study examined the fluctuation in the monthly interest rate of two particular mutual funds (i.e. Fund A and Fund B) offered by the Bank of Montreal. The following table gives the end-of-month interest rates of both of these mutual funds for 7 randomly selected months in 2009:

Fund A (x)	7.31	10.50	5.86	9.25	13.75	7.87	7.22
Fund B (y)	3.17	3.46	3.05	3.30	3.50	3.25	3.12

(4 marks) (a) Calculate the median and IQR of the Fund A data.

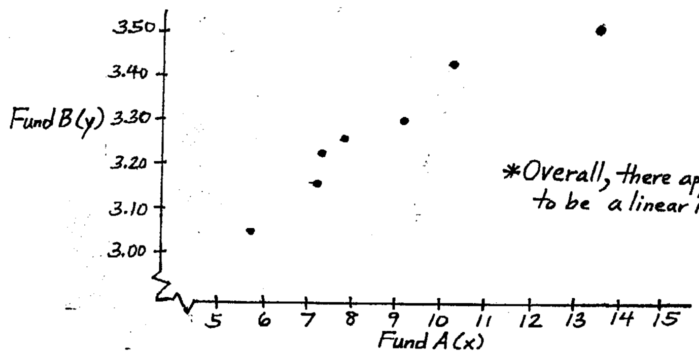
First, order the observations:

5.86 7.22 7.31 7.87 9.25 10.50 13.75

$$\tilde{x} = \text{median} = 7.87$$

$$IQR = Q_3 - Q_1 = \frac{9.25+10.50}{2} - \frac{7.22+7.31}{2} = 9.875 - 7.265 = 2.61$$

(4 marks) (b) Construct a scatter plot of the above data. Does there appear to be a linear relationship between the interest rates of these two funds?



*Overall, there appears to be a linear relationship.

(4 marks) (c) Compute the correlation coefficient and interpret it in the context of the question.

We need to compute: $\sum_{i=1}^{7} x_i = 7.31 + \dots + 7.22 = 61.76$

$$\sum_{i=1}^{7} y_i = 3.17 + \dots + 3.12 = 22.85$$

$$\sum_{i=1}^{7} x_i y_i = (7.31)(3.17) + \dots + (7.22)(3.12) = 204.1296$$

$$\sum_{i=1}^{7} x_i^2 = (7.31)^2 + \dots + (7.22)^2 = 586.716$$

$$\sum_{i=1}^{7} y_i^2 = (3.17)^2 + \dots + (3.12)^2 = 74.7599$$

Then, $s_x = \sqrt{\frac{1}{6}\left[586.716 - \frac{(61.76)^2}{7}\right]} = 2.640$

$s_y = \sqrt{\frac{1}{6}\left[74.7599 - \frac{(22.85)^2}{7}\right]} = 0.169$

Thus, $Cor(x,y) = \dfrac{\frac{1}{6}\left[204.1296 - \frac{(61.76)(22.85)}{7}\right]}{(2.640)(0.169)} = 0.945$

*The correlation is very close to +1, indicating ~~that~~ there is strong evidence of a POSITIVE linear association between the 2 interest rates.

2. The annual rates of return on 12 randomly sampled foreign stocks are as follows:

8.8 3.3 4.6 2.9 9.3 7.1 5.7 5.5 5.2 6.6 5.2 6.1

[3 marks] (a) Construct an appropriate stem and leaf plot for the above data.

First, order the observations:

2.9 3.3 4.6 5.2 5.2 5.5
5.7 6.1 6.6 7.1 8.8 9.3

Stem and Leaf Plot:

Stem = ones digit
Leaf = tenths digit

```
2 | 9
3 | 3
4 | 6
5 | 2 2 5 7
6 | 1 6
7 | 1
8 | 8
9 | 3
```

[3 marks] (b) Calculate (to 3 decimal places of accuracy) the mean and standard deviation of this data.

$$\bar{x} = \frac{1}{12}(8.8+3.3+\ldots+6.1) = \frac{70.3}{12} = 5.858$$

$$s = \sqrt{\frac{1}{12-1}\left[(8.8^2+3.3^2+\ldots+6.1^2) - \frac{(70.3)^2}{12}\right]}$$
$$= \sqrt{\frac{1}{11}(452.39 - 411.8408)}$$
$$= 1.920$$

[4 marks] (c) Does this data conform to the Empirical Rule? Justify your answer.

Look at the intervals: $(\bar{x}-s, \bar{x}+s) = (5.858-1.920, 5.858+1.920)$
$= (3.938, 7.778)$

$8/12 = 66.7\%$ of the data lie in this interval

$(\bar{x}-2s, \bar{x}+2s) = (5.858-3.840, 5.858+3.840)$
$= (2.018, 9.698)$

$12/12 = 100\%$ of the data lie in this interval

* Empirical Rule says approx. 68% and 95% should lie in these intervals $\Longrightarrow$ Based on our findings, data does not quite conform.