Key Problem Set #1

24 a) $\bar{x} = \dfrac{3+5+4+6+10+5+6+9+2+8}{10} = 5.8$

b) $2\ 3\ 4\ 5\ \underbrace{5\ 6}_{m}\ 6\ 8\ 9\ 10$ $\quad m = 5.5$

c) mode from part b we can see is 5 and 6

---

30 a) $\bar{x} = \dfrac{10+9+13+8+9+10+10+6+5+11}{10} = 9.1$

b) $5\ 6\ 8\ 9\ \underbrace{9\ 10}_{m}\ 10\ 10\ 11\ 13$ $\quad m = 9.5$

c) $\bar{x}$ and m are similar enough that it doesn't really matter what you use to quality the central tendencies

---

34 a) $\bar{x} = \dfrac{2+4+0+3+1}{5} = 2$

b) $s^2 = \dfrac{(2-2)^2 + (4-2)^2 + (0-2)^2 + (3-2)^2 + (1-2)^2}{5-1} = 2.5$ $\quad s = \sqrt{2.5} \approx 1.58$

c) $CV = \dfrac{1.58}{2} \times 100 = 79\%$

35 $\bar{x} = 36 \quad s = 3$

a) 68% will fall between 33 and 39
95% of observations will fall between 30 and 42
almost all will fall between 27 and 45

b) we could only say that at least 0% of the obs fall between 33 and 39
at least $1 - (1/2^2) = 3/4$ of the obs will fall between 30 and 42
at least $1 - (1/3^2) = 8/9$ of the obs will fall between 27 and 45

44 b) by putting into STATA (using the summarize command)

$\bar{x} = 26.23 \qquad s = 26.22$

c) $26.23 \pm 26.22 = (.01 \text{ to } 52.45)$

44 out of 50 lie in this interval 88%

$26.23 \pm 2*(26.22) = (-26.21 \text{ to } 78.67)$

47 out of 50 lie in this interval 94%

$26.23 \pm 3*(26.22) = (-52.43 \text{ to } 104.89)$

49 out of 50 lie in this interval 98%

- this matches with Tchebysheff's theorem were the % needed to be at least 0% 75% and 89% respecivly

- this does not match the emperical rule were we would expect 68%, 95%, almost all respectivly

53 a) $\bar{x} = \dfrac{3+9+6+5+5+4+7+6+8+2+6+7+3}{13} \approx 5.46$

$s^2 = \dfrac{(3-\bar{x})^2+(9-\bar{x})^2+(6-\bar{x})^2+2(5-\bar{x})^2+(4-\bar{x})^2+2(7-\bar{x})^2+2(6-\bar{x})^2+(8-\bar{x})^2+(2-\bar{x})^2+(3-\bar{x})^2}{12}$

≈ 4.269 so $s \approx \sqrt{4.269} \approx 2.066$

b) $z_{smallest} = \dfrac{2-5.46}{2.066} = -1.67$ - we are within 2 s.d. of the mean so this is reasonable

$z_{largest} = \dfrac{9-5.46}{2.066} = 1.713$ - still within 2 s.d. of the mean so reasonable

58 $\bar{x} = 410 \qquad s = 14$

$z_{score} = \dfrac{430-410}{14} = 1.43$ - this is not unusually high becouse it is still within 2 s.d of the mean

77 a) $\bar{x}_{13} = \dfrac{10+9+13+8+9+10+10+6+5+11}{10} = 9.1$

$s^2_{13} = \dfrac{.9^2+.1^2+3.9^2+1.1^2+.1^2+.9^2+.9^2+3.1^2+4.1^2+1.9^2}{9} = 5.433 \quad s = 2.33$

$\bar{x}_{30} = \dfrac{28+21+28+26+17+12+15+9+20+18}{10} = 19.4$

$s^2_{30} = \dfrac{8.6^2+1.6^2+8.6^2+6.6^2+2.4^2+7.4^2+4.4^2+10.4^2+.6^2+.6^2}{9} = 42.711 \quad s = 6.535$

b) $CV_{13} = \left(\dfrac{2.33}{9.1}\right) \times 100 = 25.6\%$

$CV_{30} = \left(\dfrac{6.535}{19.4}\right) \times 100 = 33.7\%$

c) the 30 gallon bags are more variable

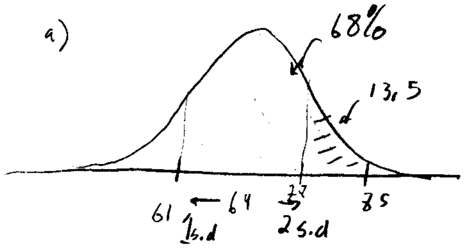
84. $\bar{x} = 69 \quad s = 8$

a) notice that 53 and 85 are 2 standard deviations (16) away from the mean so by Tchebysheff we know at least 75% of the CD prices will be between 53 and 85

b) if at least 3/4 are in the interval then at most 1/4 are outside the interval - that could be split less that 53 and more than 85 but we can't say how - so the only thing we can say is that at most 1/4 of the games will be more than $85

85 $\bar{x} = 69 \quad s = 8$ - we can now use general rule of thumb

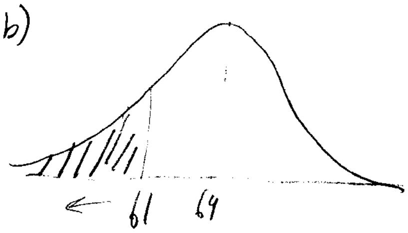
a)



adding a deviation adds $95\% - 68\% = 27\%$ that has to be split between left and right so each addition is $27/2 = 13.5\%$

- so $68\% + 13.5\% = 81.5\%$ will be between 61 and 85

b)



$100\% - 68\% = 32\%$ must be split between both sides

$32/2 = 16\%$

so 16% will be less than $61