CEE 110
Homework #7 Solution

1. The tables below delineate all 16 possible (x_1, x_2) pairs, their probabilities, the value of x for that pair, and the value of r for that pair. Probabilities are calculated using the independence of X_1 and X_2.

(x_1, x_2)	1,1	1,2	1,3	1,4	2,1	2,2	2,3	2,4
probability	.16	.12	.08	.04	.12	.09	.06	.03
$\bar{x}$	1	1.5	2	2.5	1.5	2	2.5	3
r	0	1	2	3	1	0	1	2
(x_1, x_2)	3,1	3,2	3,3	3,4	4,1	4,2	4,3	4,4
probability	.08	.06	.04	.02	.04	.03	.02	.01
$\bar{x}$	2	2.5	3	3.5	2.5	3	3.5	4
r	2	1	0	1	3	2	1	0

a. Collecting the $\bar{x}$ values from the table above yields the pmf table below.

$\bar{x}$	1	1.5	2	2.5	3	3.5	4
P($\bar{x}$)	.16	.24	.25	.20	.10	.04	.01

b. $P(\bar{x} \le 2.5) = .16 + .24 + .25 + .20 = .85$.

c. With n = 4, there are numerous ways to get a sample average of at most 1.5, since $\bar{x} \le 1.5$ if the sum of the X_i is at most 6. Listing out all options, $P(\bar{x} \le 1.5) = P(1,1,1,1) + P(2,1,1,1) + \cdots + P(1,1,1,2) + P(1,1,2,2) + \cdots + P(2,2,1,1) + P(3,1,1,1) + \cdots + P(1,1,1,3) = (.4)^4 + 4(.4)^3(.3) + 6(.4)^2(.3)^2 + 4(.4)^2(.2)^2 = .2400$

d. Collecting the r values from the table above yields the pmf table below.

r	0	1	2	3
$p(r)$	.30	.40	.22	.08

2. μ = 70 GPa, σ = 1.6 GPa

a. The sampling distribution of $\bar{X}$ is centered at $E(\bar{X}) = \mu = 70$ GPa, and the standard deviation of the $\bar{X}$ distribution is $\sigma_{\bar{x}} = \frac{\sigma_{\bar{x}}}{\sqrt{n}} = \frac{1.6}{\sqrt{16}} = 0.4$ GPa.

b. With n = 64, the sampling distribution of $\bar{X}$ is still centered at $E(\bar{X}) = \mu = 70$ GPa, but the standard deviation of the $\bar{X}$ distribution is $\sigma_{\bar{x}} = \frac{\sigma_{\bar{x}}}{\sqrt{n}} = \frac{1.6}{\sqrt{64}} = 0.2$ GPa.

c. $\bar{X}$ is more likely to be within 1 GPa of the mean (70 GPa) with the second, larger, sample. This is due to the decreased variability of $\bar{X}$ that comes with a larger sample size.

d. In the previous exercise, we found $E(\bar{X}) = 70$ and $SD(\bar{X}) = 0.4$ when n = 16. If the diameter distribution is normal, then $\bar{X}$ is also normal, so $P(69 \leq \bar{X} \leq 71) = P\left(\frac{69-70}{0.4} \leq Z \leq \frac{71-70}{0.4}\right) = P(-2.5 \leq Z \leq 2.5) = \Phi(2.5) - \Phi(-2.5) = .9938 - .0062 = .9876.$

e. With n = 25, $E(\bar{X}) = 70$ but $SD(\bar{X}) = \frac{1.6}{\sqrt{25}} = 0.32$ GPa. So, $P(\bar{X} \geq 71) = P\left(Z > \frac{71-70}{0.32}\right) = 1 - \Phi(3.125) = 1 - .9991 = .0009.$

3.

a. $58.3 \pm \frac{1.96(3)}{\sqrt{25}} = 58.3 \pm 1.18 = (57.1, 59.5).$

b. $58.3 \pm \frac{1.96(3)}{\sqrt{100}} = 58.3 \pm .69 = (57.7, 58.9).$

c. $58.3 \pm \frac{2.58(3)}{\sqrt{100}} = 58.3 \pm .77 = (57.5, 59.1).$

d. $n = \left[\frac{2(2.58)3}{1}\right]^2 = 239.63 \approx 240.$

4.

a. From the data provided, $\bar{x} = 107.78$ and $s = 1.076$. The corresponding 95% CI for μ is $\bar{x} \pm t_{0.025,5-1} \frac{s}{\sqrt{n}} = 107.78 \pm 2.776 \frac{1.076}{\sqrt{5}} = (106.44, 109.12).$

b. The CI suggests that while 107 is a plausible value for μ (since it lies in the interval).

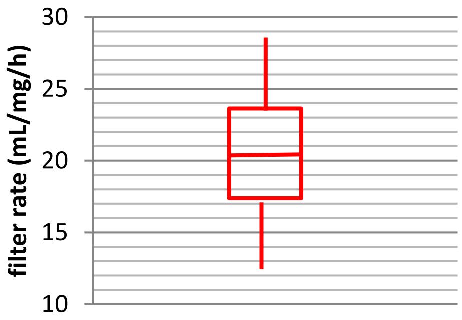
c. The corresponding 95% upper bound for μ is $\mu < \bar{x} + t_{0.05,5-1} \cdot \frac{s}{\sqrt{n}} = 107.78 + 2.132 \frac{1.076}{\sqrt{5}}$. Therefore, $\mu < 108.81$

d. The CI suggests that 110 is not a plausible value for μ (since it does not lie in the interval).

5.

a. Median = 20.54
   Q1 = (16.03 +18.45)/2=17.24
   Q3 = (22.63+25.05)/2=23.84
   Given Median, Q1 and Q3, it is normal.

   Alternatively, draw the box plot.
   IQR = 23.84-17.24= 6.6
   Min = Q1 – 1.5×IQR = 17.24-1.5×6.6 = 7.34 → 12.19
   Max= Q3 + 1.5×IQR = 23.84+1.5×6.6 = 33.74 → 28.8



   From the boxplot, it seems to be normal.

b. E(X)= sum/n = 143.8/7 = 20.54

   s =√(Σ(x-20.54)^2/(7-1))= 5.61

c. 95% interval → α = 0.05
   two-sided, as population follows normal distribution

   n = 7, $\overline{X} = 20.54$, s = 5.61; $t_{0.025,6}$ = 2.447 (from t distribution table)

$$20.54 \pm t_{.025,6}\left(\frac{5.61}{\sqrt{7}}\right) = 20.54 \pm (2.447)\left(\frac{5.61}{\sqrt{7}}\right) = 20.54 \pm 5.19$$

$$= (15.35, 25.73)$$

(or $20.54 \pm 5.189, (15.351, 25.729)$)

   99% interval → α = 0.01

   n = 7, $\overline{X} = 20.54$, s = 5.61; $t_{0.005,6}$ = 3.707 (from t distribution table)

$$20.54 \pm t_{.005,6}\left(\frac{5.61}{\sqrt{7}}\right) = 20.54 \pm (3.707)\left(\frac{5.61}{\sqrt{7}}\right) = 20.54 \pm 7.86$$

$$= (12.68, 28.4)$$

   99% confidence level provides wider interval than 95% level.

6.

a. H_0: $\mu \leq 30$ years
   H_a: $\mu > 30$ years

b. One sided, α=0.01, n=15
   The rejection region for 1% is $t_{0.01,\ 15-1} = 2.624$
   Therefore, rejection region ≥ 2.624

c. $\mu_0 = 30$
   n = 15
   s = 0.2
   $\bar{X} = 30.15$

$$t = \frac{\bar{x}-\mu_0}{s/\sqrt{n}} = \frac{30.15-30}{0.2/\sqrt{15}} = 2.905$$

t = 2.905>2.624.

Therefore, reject the null hypothesis. The half-life of cesium-137 is greater than 30 years.

d. The p-value is P(X>2.905)
   From the t curve tail areas table, P(t=2.9 ν=14) = 0.005 (or P(t=3.0 ν=14) = 0.004)

   The p-value, 0.005 < α=0.01, (less than 1%), which is consistent with the hypothesis testing. Therefore, this rejects the null hypothesis.

e. 99% interval $\rightarrow$ $\alpha = 0.01$
   n = 15, $\bar{X} = 30.15$, s = 0.2; $t_{0.01,15-1} = 2.624$ (from t distribution table)
   Lower confident bound for μ

$$30.15 - t_{0.01,14}\left(\frac{0.2}{\sqrt{15}}\right) = 30.15 - (2.624)\left(\frac{0.2}{\sqrt{15}}\right) = 30.15 - 0.14 = 30.01$$

Therefore, μ>30.01

$\mu \leq 30$ is not contained within this CI, which is consistent with the hypothesis testing problem having a p-value of 0.004, so that the null hypothesis is rejected at size α=0.01