Midterm 2 Solution

1.

a. The marginal pmf of X is

$$p_X(x) = \begin{cases} 0.30 & x = 0 \\ 0.20 & x = 1 \\ 0.50 & x = 2 \\ 0 & \text{otherwise} \end{cases}$$

The marginal pmf of Y is

$$P_Y(y) = \begin{cases} 0.26 & x = 0 \\ 0.34 & x = 1 \\ 0.20 & x = 2 \\ 0.20 & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

b. No. Because $p(0,0) = .12 \neq p_X(0)p_Y(0)$

c. The conditional pmf of $Y \mid X = 1$ is

$$p_{Y|X=1}(y) = \begin{cases} 0.2 & x = 0 \\ 0.4 & x = 1 \\ 0.3 & x = 2 \\ 0.1 & x = 3 \\ 0 & \text{otherwise} \end{cases}$$

d. $E(\bar{X}) = E(X)$ and $E(\bar{Y}) = E(Y)$. And $E(\bar{X} + \bar{Y}) = E(\bar{X}) + E(\bar{Y})$
$= E(X) + E(Y) = (0 * 0.30 + 1 * 0.20 + 2 * 0.50) + (0 * 0.26 + 1 * 0.34 + 2 * 0.20 + 3 * 0.20)$
$= 2.54$

2.

a. The joint pdf is

$$f(x_1, x_2, \ldots, x_n \mid p) = \prod_{i=1}^{n} (1-p)^{x_i - 1} p = (1-p)^{\sum x_i - n} p^n$$

The log likelihood function is

$$\text{log} f(p) = \left(\sum x_i - n\right) \text{log}(1-p) + nlogp$$

Take the derivative of the log likelihood function

$$\frac{d\text{log} f(p)}{dp} = -\frac{\left(\sum x_i - n\right)}{1-p} + \frac{n}{p}$$

Let

$$-\frac{(\sum x_i - n)}{1-p} + \frac{n}{p} = 0$$

Solve the above equation for p, we have

$$p = \frac{n}{\sum x_i}$$

Hence the maximum likelihood estimator of p is

$$\hat{p} = \frac{n}{\sum X_i}$$

Based on the observed data, the maximum likelihood estimate is $\hat{p} = 5/(2+4+2+5+1) = 5/14 = 35.7\%$

b.

$$p_3 = P(X=3) = (1-p)^{3-1}p = (1-p)^2 p$$

Applying the invariance principle of MLE, the MLE of p_3 is

$$\hat{p_3} = (1-\hat{p})^2\hat{p}$$

Based on the observed data, $\hat{p_3} = (1-.357)^2.357 = .148$

3.

$\bar{x} = .5060$, $s = .004$

a.$t_{\alpha/2,n-1} = t_{.025,9} = 2.262$, the confidence interval is

$$\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}} = .506 \pm 2.262\frac{.004}{\sqrt{10}} = .506 \pm .0029$$

b. In probability, 95% of all CIs constructed with $\bar{x} \pm t_{\alpha/2,n-1}\frac{s}{\sqrt{n}}$ contains the actual average diameter. So, we are confident that $(.506 \pm .0029)$ covers the true parameter (it is one of those CIs).

c. $H_0 : \mu = 0.5$ vs $H_a : \mu \neq 0.5$

$$t_{obs} = \frac{\bar{x} - 0.5}{s/\sqrt{n}} = \frac{.5060 - .5}{.004/\sqrt{10}} = 4.74$$

Since t_{obs} is in the rejection region $(-\infty, -2.262) \bigcup (2.262, \infty)$, H_0 is rejected.

4. $H_0 : p = 0.5$ vs $H_a : p > 0.5$

Test statistic x=number of successes in 20 attempts=13. Since $\alpha = 5\%$, and

$$1 - B(13; 0.5, 20) = 1 - .942 = .058$$

$$1 - B(14; 0.5, 20) = 1 - .979 = .021$$

The rejection region is $R = \{15, 16, \ldots, 20\}$
Accept H_0 because $x = 13$ is not in the rejection region.

5.
a. $H_0 : \mu = 4.5$ vs $H_a : \mu < 4.5$

$$z_{obs} = \frac{\bar{x} - \mu_0}{\sigma/\sqrt{n}} = \frac{3.9 - 4.5}{1.5/\sqrt{25}} = -2$$

The rejection region is $(-\infty, -1.645)$. Because z_{obs} is in the rejection region, H_0 is rejected.
b. P-value=$P(Z \leq z_{obs} \mid H_0)$=$P(Z \leq -2 \mid H_0)$ =.0228
c.
$P(\text{type I}) = \alpha = 5\%$
Rejection region for $\bar{x}$:

$$(-\infty, \mu_0 - 1.645\frac{\sigma}{\sqrt{n}}) = (-\infty, 4.5 - 1.645\frac{1.5}{5}) = (-\infty, 4.0065)$$

Acceptance region is $(4.0065, +\infty)$

$$P(\text{type II}) = P(\bar{X} \geq 4.0065 \mid \mu' = 4.0) = P(\frac{\bar{x} - 4.0}{\sigma/\sqrt{n}} \geq \frac{4.0065 - 4.0}{1.5/\sqrt{25}})$$
$$= P(Z \geq 0.02167) = .4129$$

d.

$$n = \left[\frac{\sigma(z_\alpha + z_\beta)}{\mu - \mu'}\right]^2 = \left[\frac{1.5(1.645 + 1.645)}{4.5 - 4.0}\right]^2$$
$$= (3 * 1.645 * 2)^2 = 97.4$$

Hence, the sample size should be 98.