Q1. $P(X \leq Y) = \frac{1}{8}\int_0^2 \int_0^y x + y\, dx\, dy$

$= \frac{1}{8}\int_0^2 \left[\frac{1}{2}x^2 + xy\right]_0^y dy$

$= \frac{1}{8}\int_0^2 \frac{1}{2}y^2 + y^2\, dy = \frac{1}{8}\left(\frac{3}{2}\right)\int_0^2 y^2\, dy$

$= \frac{3}{16}\left[\frac{1}{3}y^3\right]_0^2$

$= \frac{3}{16}\left(\frac{8}{3}\right) = \frac{1}{2}$

Q2. $Let\ X = number\ of\ websites\ needed\ to\ visit\ in\ order\ to\ get\ the\ first\ occurence\ of\ a\ keyword$

$X \sim Geo(p = 0.3)$

$P(X \geq 4) = 1 - P(X \leq 3)$

$= 1 - [0.3 + 0.3(0.7) + 0.3(0.7^2)]$

$= 1 - 0.657$

$= 0.343$

Q3. (a) $Let\ X = number\ of\ girls\ in\ a\ family$

$X \sim Bin(n = 5, p = 0.5)$

$P(X \geq 1) = 0.9688$

(b) $P(X \geq 1 | X \leq 4) = \frac{P(1 \leq X \leq 4)}{P(X \leq 4)} = \frac{P(X \geq 1) - P(X \geq 5)}{1 - P(X \geq 5)}$

$= \frac{0.9688 - 0.0313}{1 - 0.0313} = \frac{0.9375}{0.9687} = 0.9678$

Q4. (a) $Let\ X =\ number\ of\ train\ delays\ in\ a\ year$

$X \sim Poi(\lambda = 10), \lambda\ in\ year$

$P(X > 10) = P(X \geq 11) = 0.4170$

(b) $Let\ Y =\ waiting\ time\ in\ year\ until\ the\ next\ delay\ appears$

$Y \sim exp(\lambda = 10), \lambda\ in\ year$

$F(y) = 1 - e^{-10y}$

$3\ months \rightarrow\ \frac{3}{12}\ or\ 0.25\ year$

$P(Y \leq 0.25) = F(0.25) = 1 - e^{-10(0.25)} = 1 - e^{-2.5} = 0.9179$

(c) $Let\ T =\ waiting\ time\ in\ year\ until\ the\ third\ delay\ appears$

$T \sim \Gamma(\alpha = 3, \lambda = 10), \lambda\ in\ year$

$6\ months \rightarrow\ t = 0.5\ year$

$Using\ Gamma - Poisson\ formula,$

$X \sim Poi(\lambda t = 5),\ \ \lambda t = 10(0.5) = 5$

$P(T < 0.5) = P(X \geq 3) = 0.8753$

Q5. (a) $\hat{p}_X = \frac{15}{300} = 0.05$

$\hat{p}_Y = \frac{8}{200} = 0.04$

$\therefore The\ point\ estimate\ of\ p_X - p_Y\ is\ 0.05 - 0.04 = 0.01$

(b) $n = 300,\ \ \hat{p}_X = 0.05,\ \ \hat{q}_X = 0.95$

$m = 200,\ \ \hat{p}_Y = 0.04,\ \ \hat{q}_Y = 0.96$

$n\hat{p}_X = 15,\ \ n\hat{q}_X = 285,\ \ m\hat{p}_Y = 8,\ \ m\hat{q}_Y = 192$

$Since\ these\ values\ are\ all\ greater\ than\ 5, we\ will\ use\ z\ test\ to\ make\ a\ confidence\ interval\ for\ p_X - p_Y.$

$\alpha = 0.02, z_{\frac{\alpha}{2}} = 2.3263$

$0.01 \pm 2.3263\sqrt{\frac{0.05(0.95)}{300} + \frac{0.04(0.96)}{200}} = 0.01 \pm 0.0435$

$\therefore A\ 98\%\ CI\ for\ p_X - p_Y\ is\ (-0.0335, 0.0535).$

Q6. (a)

Here, σ is not known, $n < 30$ and the population is normally distributed. Therefore, we will use the *t* distribution to make a confidence interval for μ.

(b)

$n = 28\ (< 30),\ \ \bar{x} = 24.1279,\ \ s = 16.8322\ (\sigma\ not\ known),\ \ s_{\bar{x}} = 3.1810$

$\alpha = 0.05, \nu = 27, t_{\frac{\alpha}{2},27} = 2.052$

$24.1279 \pm 2.052(3.1810) = 24.1279 \pm 6.5274$

$\therefore A\ 95\%\ CI\ for\ \mu\ is\ (17.6005, 30.6553).$

We are 95% confident that the average amount of money spent per week at coffee shops by all customers who visit coffee shops in Sungai Long area is somewhere between RM 17.60 and RM 30.66.