ASSIGNMENT- 4

10.7 a) Observed and expected counts for first school are

obs (i, j)	Pass	Fail	Total
Girls	162	69	231
Boys	567	376	945
Total	729	447	1176

$$\text{Exp}(1,1) = \frac{729 \times 231}{1176} \qquad \text{Exp}(1,2) = \frac{447 \times 231}{1176}$$

$$\text{Exp}(2,1) = \frac{947 \times 231}{1176} \qquad \text{Exp}(2,2) = \frac{447 \times 945}{1176}$$

Exp(i, j)	Pass	Fail	Total
Girls	143.196	87.804	231
Boys	585.804	359.196	945
Total	729	447	1176

$$X^2_{obs} = \sum_{k=1}^{N} \left\{ \frac{obs(k) - \text{Exp}(k)}{\text{Exp}(k)} \right\}^2$$

$$= \frac{(162 - 143.196)^2}{143.196} + \frac{(69 - 87.804)^2}{87.804} + \cdots = 8.084$$

$$P(X^2_{obs} \geq 8.084) \rightarrow \text{b/w } 0.001 \text{ \& } 0.005$$

There is significant evidence that girls and boys perform differently as the AP test in the 1st school b)

obs(i, j)	Pass	Fail	Total
Girls	462	693	1155
Boys	57	132	189
Total	519	825	1344

d

$Exp(1,1) : \frac{519 \times 1155}{1344}$ $Exp(1,2) : \frac{519 \times 189}{1344}$

$Exp(2,1) : \frac{825 \times 1155}{1344}$ $Exp(2,2) = \frac{825 \times 184}{1344}$

$\chi^2_{obs} \cdot \frac{(Obs(k) - Exp(k))^2}{Exp(k)}$

$\frac{(462 - 446.016)^2}{446.016} + \frac{(693 - 708.894)^2}{708.844} + \dots = 6.636$

$P(\chi^2_{obs} \geq 6.636) = $ value between 0.005 & 0.01

There is significant evidence that girls and boys perform differently on the AP test in the 2nd school

c)

obs(i, j)	Pass	Fail	Total
Girls	624	762	1386
Boys	624	510	1134
Total	1248	1272	2520

$Exp(1,1) : \frac{1248 \times 1386}{2520}$ $Exp(1,2) : \frac{1248 \times 1134}{2520}$

$Exp(2,1) : \frac{1272 \times 1386}{2520}$ $Exp(2,2) \cdot \frac{1272 \times 1134}{2320}$

Exp(i, j)	Pass	Fail	Total
Girls	686.4	699.6	1386
Boys	561.6	572.4	1134
Total	1248	1272	2520

$\chi^2_{obs} : \frac{(obs(k) - Exp(k))^2}{Exp(k)}$

$$\frac{(624-686.4)^2}{686.4} + \frac{(762-699.6)^2}{699.6} \cdots = 24.974$$

$P(X^2 \geq 24.974) \rightarrow$ value < 0.001

There is significant evidence that girls and boys perform differently as the AP test in the 2 schools.

10.31 $X = 4$ accidents / month

Poisson (θ) distribution

prior distribution $\rightarrow$ Gamma$(5, 1)$

$\alpha = 5, \lambda = 1$

$X \sim Poi(\theta)$

$$P(4/\theta) = \frac{e^{-\theta}\theta^4}{4!} \propto e^{-\theta}\theta^4$$

$$\pi(\theta) = \frac{1^5}{\Gamma(5)} e^{-\theta}\theta^4 \propto e^{-\theta}\theta^4$$

posterior $\pi(\theta/x) \propto P(x/\theta) \cdot \pi(\theta)$

$= e^{-\theta}\theta^4 \times e^{-\theta}\theta^4$

$= e^{-2\theta}\theta^8$

$= e^{-2\theta}\theta^{9-1}$

guessing posterior as of θ given $x = 4$ is Gamma $(9, 2)$

Mean $= \frac{\alpha}{\lambda}$ $\qquad \hat{\theta}_{Bayes} = E(\theta / X=4) = \frac{9}{2} = 4.5$

Variance $= \frac{\alpha}{\lambda^2}$

posterior $\hat{\theta}_{Bayes} = E(\theta/x) \rightarrow$

$R(x) = E((\theta - E(\theta/x))^2/x) = Var(x/\theta)$

$$R(4) = Var(\theta/4) = \frac{9}{2^2} = \frac{9}{4} = 2.25$$

10.34 a) $\mu = \theta \quad \sigma = 4000$

prior of $\theta \rightarrow$ Normal $(14000, 2000)$

$$\hat{\theta}_B = \mu_x = \frac{\frac{n\bar{X}}{\sigma^2} + \frac{\mu}{T^2}}{\frac{n}{\sigma^2} + \frac{1}{T^2}} = \frac{\frac{50 \times 17.95}{(3.16)^2} + \frac{14000}{(2000)^2}}{\frac{50}{(3.16)^2} + \frac{1}{(2000)^2}}$$

$= 17.661$ (In thousands)

b) $1 - \alpha = 0.9$

$\alpha = 0.1 \quad \frac{\alpha}{2} = 0.05$

$$T_x = \frac{1}{\sqrt{\frac{n}{\sigma^2} + \frac{1}{T^2}}} = \frac{1}{\sqrt{\frac{50}{(3.16)^2} + \frac{1}{(2000)^2}}} = 0.4468$$

$\mu_x \pm Z_{0.05} T_x$

Set $\rightarrow [17.661 - (1.64)(0.4468), 17.661 + (1.64)(0.4468)]$

---- $[16.765, 18.556]$ in thousands of users

Given the observed data & the assumed prior distribution there is a 90% posterior probability that the average number of concurrent users is between 16765 and 18556

c) Yes there is significant evidence as the 90% credible set contains the value range way over 16000

10.36 a) 2.5, 7.4, 8, 4.5, 7.4, 9.2

prior → [5.0, 6.0] with 0.95, $1-\alpha = 0.95$

$\alpha = 0.05$

prior mean $\mu = \frac{6+5}{2} = 5.5$

pri std = 2.2

$\tau = \frac{6-5}{2 \times Z_{0.025}} = \frac{1}{2 \times 1.96} = 0.255$

$\pi(\theta)$: Normal (5.5, 0.255)

b) $\bar{x} = 6.5$

$n = 6$

$$\mu_x = \frac{\frac{n\bar{x}}{\sigma^2} + \frac{\mu}{\tau^2}}{\frac{n}{\sigma^2} + \frac{1}{\tau^2}} = \frac{\frac{6 \times 6.5}{(2.2)^2} + \frac{5.5}{(0.255)^2}}{\frac{6}{(2.2)^2} + \frac{1}{(0.255)^2}} = 5.575$$

$$\tau_x = \frac{1}{\sqrt{\frac{n}{\sigma^2} + \frac{1}{\tau^2}}} = \frac{1}{\sqrt{\frac{6}{(2.2)^2} + \frac{1}{(0.255)^2}}} = 0.2453$$

posterior $= \tau_x^2$

$= 0.06$

c) 95% HPD set would be

$\mu_x \pm Z_{0.025}\, \tau_x$

$[5.75 - (1.96)\, 0.2453),\ 5.575 + (1.96)(0.2453)]$

$[5.0842,\ 6.0337]$

95% HPD would be much more narrow because besides the data, it uses HPD a informative (low variance) prior distribution