Solutions to 2014 MAS exam: Theory

1.  
   a) 
   \[ P(\text{gray} | \text{litter 2}) = \frac{2}{5}. \]
   
   b) 
   \[ P(\text{brown}) = P(\text{brown} | \text{litter 1})P(\text{litter 1}) + P(\text{brown} | \text{litter 2})P(\text{litter 2}) \]
   \[ = \left(\frac{2}{3}\right)\left(\frac{1}{2}\right) + \left(\frac{3}{5}\right)\left(\frac{1}{2}\right) = \frac{19}{30}. \]
   
   c) 
   \[ P(\text{litter 1} | \text{brown}) = \frac{P(\text{brown} | \text{litter 1})P(\text{litter 1})}{P(\text{brown})} \]
   \[ = \frac{\left(\frac{2}{3}\right)\left(\frac{1}{2}\right)}{\frac{19}{30}} = \frac{10}{19}. \]

2.  
   A negative binomial random variable has mean \( \frac{r}{p} \) and variance \( \frac{r(1-p)}{p^2} \).
   
   a) By the weak law of large numbers (Mean of a sequence of independent random variables, each with mean \( \mu \) and variance \( \sigma^2 \), converges in probability to the mean \( \mu \)), we have 
   \[ \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \] converges in probability to \( \frac{r}{p} \). That means, for any given small value \( \epsilon \), the probability that \( \bar{X} \) is beyond the given distance \( \epsilon \) of \( \frac{r}{p} \) goes to zero as \( n \) increases to \( \infty \).
   
   b) Central limit theorem tells “For a sequence of independent and identically distributed random variables, each with mean \( \mu \) and variance \( \sigma^2 \), \( \frac{\bar{X}}{\sigma/\sqrt{n}} \) converges in distribution to \( N(0,1) \).” By the Central limit theorem, we have 
   \[ \frac{\sqrt{n}(\bar{X} - \frac{r}{p})}{\sqrt{\frac{r(1-p)}{p^2}}} \] converges in distribution to \( N(0,1) \).
c) For a negative binomial random variable with \( n=30, \ r=10, \ p=.5 \), mean is \( r/p = 20 \) and variance is \( r(1-p)/p^2 = 20 \).

\[
P(\bar{X} \leq 11) = P\left(\frac{\sqrt{30}(\bar{X} - 20)}{\sqrt{20}} \leq \frac{\sqrt{30}(19 - 20)}{\sqrt{20}}\right) \approx P(Z \leq -1.225) = 0.1103
\]

3.

a) The marginal density of \( X \) is

\[
f_X(x) = \int_0^1 f_{X,Y}(x,y)dy
\]
\[= \int_0^1 \frac{1}{4}(x + 2y)dy
\]
\[= \frac{1}{4}(xy + y^2)|_y^1
\]
\[= \frac{1}{4}(x + 1), \ 0 \leq x < 2.
\]

b) The conditional density of \( Y \) given \( X \) is

\[
f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}
\]
\[= \frac{\frac{1}{4}(x + 2y)}{\frac{1}{4}(x + 1)}
\]
\[= \frac{x + 2y}{x + 1}, \ 0 \leq x < 2, \ 0 \leq y < 1.
\]

\( X \) and \( Y \) are not independent because the conditional density of \( Y \) given \( X \) depends on \( X \).

c) \( Z = \frac{9}{(X+1)} \Rightarrow X = \frac{9}{Z} - 1 \). So, \( 3 < z \leq 9 \) and \( J = \frac{dx}{dz} = -\frac{9}{z^2} \). Therefore, the density of \( Z \) is

\[
f_Z(z) = \frac{1}{4}\left(\frac{9}{z} - 1 + 1\right)|J| = \frac{81}{4}z^{-3}, \ 3 < z \leq 9.
\]
4.

a) By looking at the density function, it is actually a beta distribution with parameters \(a = \theta\) and \(b = 1\). Therefore, mean of the distribution is \(\frac{\theta}{\theta+1}\). Solving \(\frac{\hat{\theta}}{\hat{\theta}+1} = \bar{X}\), we have the method of moments estimator of \(\theta\) is

\[
\hat{\theta} = \frac{\bar{X}}{1 - \bar{X}}.
\]

b) The log likelihood function of \(\theta\) is

\[
l(\theta) = \log \left( \prod_{i=1}^{n} f(x_i|\theta) \right)
= \log \left( \prod_{i=1}^{n} \theta x_i^{\theta-1} \right)
= \log(\theta^n \left( \prod_{i=1}^{n} x_i \right)^{\theta-1})
= n \log \theta + (\theta - 1) \sum_{i=1}^{n} \log x_i.
\]

Then \(\frac{\partial l(\theta)}{\partial \theta} = \frac{n}{\theta} + \sum_{i=1}^{n} \log x_i\). Solving \(\frac{\partial l(\theta)}{\partial \theta} = 0\), we have

\[
\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log x_i}.
\]

Because the second derivative of \(l(\theta)\) is \(-\frac{n}{\theta^2}\), which is always negative, the solution above is a maximizer of \(l(\theta)\). Therefore \(\hat{\theta} = \frac{n}{\sum_{i=1}^{n} \log x_i}\) is the MLE of \(\theta\).

c) By the large sample theory, we know the asymptotic variance of the MLE is the inverse of the fisher information, which is

\[
I(\theta)^{-1} = \left( -E\left[ \frac{\partial^2}{\partial \theta^2} l(\theta) \right] \right)^{-1}
= \left( -E\left( -\frac{n}{\theta^2} \right) \right)^{-1}
= \left( \frac{n}{\theta^2} \right)^{-1} = \frac{\theta^2}{n}.
\]
5.

a) The joint probability mass function of the six random variables, $X_1, \ldots, X_6$, is given by

$$f(x_1, x_2, x_3, x_4, x_5, x_6|\theta) = \frac{e^{-6\theta} \cdot \theta^{\sum_{i=1}^{6} x_i}}{\prod_{i=1}^{6} x_i!}.$$ 

The prior distribution of $\theta$ is given by $\text{Gamma}(2, 1.5)$,

$$\pi(\theta) = \frac{\Gamma^2}{1.5^2} \theta^{2-1} e^{-1.5\theta}.$$ 

Thus for the posterior distribution we have the following:

$$\pi(\theta|x) \propto f(x|\theta) \cdot \pi(\theta)$$

$$\propto e^{-6\theta} \cdot \theta^{\sum_{i=1}^{6} x_i} \cdot \theta^{2-1} e^{-1.5\theta}$$

$$\propto \theta^{\sum_{i=1}^{6} x_i+2-1} e^{-7.5\theta},$$

which is the kernel of the $\text{Gamma}\left(\left[\sum_{i=1}^{6} x_i + 2\right], 7.5\right)$ distribution. Therefore, the posterior distribution of $\theta$ is $\text{Gamma}\left(\left[\sum_{i=1}^{6} x_i + 2\right], 7.5\right)$.

b) A Bayes estimator is the mean of the posterior distribution. Therefore, a Bayes estimator of $\theta$ can be

$$\hat{\theta} = \frac{\sum_{i=1}^{6} X_i + 2}{7.5}$$

c) A 90% credible interval for $\theta$ can be given by the 5th and the 95th percentile of the $\text{Gamma}\left(\left[\sum_{i=1}^{6} x_i + 2\right], 7.5\right)$ distribution.
6.

a) The common probability density function of $X_1, X_2, \ldots, X_5$ is given by

$$f(x) = \theta e^{-\theta x}, \ x \geq 0.$$  

To derive the distribution of $Y$, we calculate

$$P(Y > y) = P(X_1 > y, X_2 > y, \ldots, X_5 > y)$$

$$= [P(X_i > y)]^5$$

$$= \left[ \int_y^\infty \theta e^{-\theta x} \, dx \right]^5 = e^{-5\theta y}$$

Hence $f_Y(y) = 5\theta e^{-5\theta y}$.

b) The power function of the test is given as follows:

$$\pi(\theta) = P_\theta(\text{rejecting } H_0)$$

$$= P_\theta(Y > 18)$$

$$= e^{-5\theta \cdot 18} = e^{-90\theta}$$

Note that

$$\theta_1 > \theta_2 \Rightarrow -90\theta_1 < -90\theta_2$$

$$\Rightarrow e^{-90\theta_1} < e^{-90\theta_2}$$

Therefore the power function is a decreasing function of $\theta$.

c) The size (the significance level) of the test is:

$$\sup_{\theta \in \Omega_0} \pi(\theta) = \sup_{\theta \geq \frac{1}{20}} e^{-90\theta} = e^{-90/20} = 0.011109$$