Assignment 2 Chunlei Zhou

Q1.

(a) The density of binomial distribution is

$$P(X = x\,|p) = \binom{x-1}{r-1} p^r (1-P)^{x-r}$$

The log-likelihood is

$$L(p; r) = rlog(p) + (x-r)\log(1-p) + C$$

C is a constant.
When

$$\frac{dL(p;r)}{dp} = \frac{r}{p} - \frac{x-r}{1-p} = 0$$

We can have

$$\hat{p}_{MLE} = \frac{r}{X}$$

Since

$$E[X] = \mu = \frac{r}{p}$$

So

$$\hat{\mu}_{MLE} = \frac{r}{\hat{p}MLE} = \frac{r}{(\frac{r}{X})} = X$$

(b) p ~ Beta (α, β), X ~ nb (r, p). We can write the density of the posterior distribution

$$\pi(p|x) = \frac{P(X = x\,|p)\pi(p)}{\int_{p=0}^{1} P(X = x\,|p)\pi(p)dp}$$

$$= \frac{\binom{x-1}{r-1} p^r (1-p)^{x-r} \frac{1}{Beta(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1}}{\int_{p=0}^{1} \binom{x-1}{r-1} p^r (1-p)^{x-r} \frac{1}{Beta(\alpha,\beta)} p^{\alpha-1}(1-p)^{\beta-1} dp}$$

$$= \frac{\binom{x-1}{r-1} p^{r+\alpha-1}(1-p)^{x-r+\beta-1}}{\int_{p=0}^{1} \binom{x-1}{r-1} p^{r+\alpha-1}(1-p)^{x-r+\beta-1}\, dp}$$

We have

$$Beta(r+\alpha, x-r+\beta) = \int_{u=0}^{1} u^{r+\alpha-1}(1-u)^{x-r+\beta-1}$$

It is obvious that the beta density is a conjugate prior for p.

(c) The prior distribution of p is Beta (α, β). The density of which is

$$f(p|\alpha,\beta) = \frac{1}{Beta(\alpha,\beta)}\, p^{\alpha-1}(1-p)^{\beta-1}$$

$$\hat{p}_{prior} = \frac{\alpha}{(\alpha+\beta)}$$

$$\hat{\mu}_{prior} = r/\,\hat{p}_{prior} = \frac{r(\alpha+\beta)}{\alpha}$$

The density of posterior distribution of p is given in solution (b).
We have proved that p~ $\mathrm{Beta}(\mathrm{r}+\alpha, \mathrm{x}-\mathrm{r}+\beta)$
Thus, we can have α' = α + r; β' = x - r + β

$$\hat{p}_{post} = \frac{\alpha'}{\alpha'+\beta'} = \frac{\alpha+r}{(\alpha+r+x-r+\beta)} = \frac{\alpha+r}{\alpha+\beta+x}$$

$$\hat{\mu}_{post} = r/\,\hat{p}_{post} = \frac{r(\alpha+\beta+x)}{\alpha+r}$$

(d) Since we have

$$\hat{\mu}_{post} = \frac{r(\alpha+\beta+x)}{\alpha+r}$$

We can decompose the equation above to

$$\hat{\mu}_{post} = \frac{r(\alpha+\beta)}{\alpha+r} + \frac{xr}{\alpha+r}$$

$$= x * \frac{r}{(\alpha+r)} + \frac{r(\alpha+\beta)}{\alpha} * \frac{\alpha}{\alpha+r}$$

$$= \frac{r}{\alpha+r} * \hat{\mu}_{MLE} + \frac{\alpha}{\alpha+r} * \hat{\mu}_{prior}$$

$$= q\hat{\mu}_{MLE} + (1-q)\hat{\mu}_{prior}$$

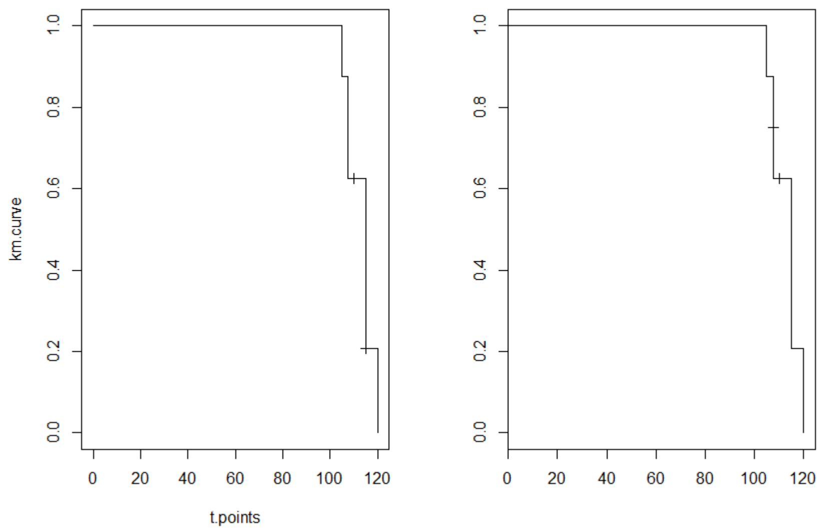
$$where\ q = \frac{r}{\alpha+r}$$

Q2.

(a)

i	ti	di	r(ti)	$\hat{p}_i$
0	0	0	8	(8-0)/8 = 1
1	105	1	8	(8-1)/8 = 7/8
2	107.5	2	7	(7-2)/7 = 5/7
3	110	0	4	(4-0)/4 = 1
4	115	2	3	(3-2)/3 = 1/3
5	120	1	1	(1-1)/1 = 0

(b)



The two plots are almost the same with slight different in censored label.

Q3.

(a)

Given the hazard rate

$$h(x) = h_0(x)e^{\eta}, x > 0$$

The cumulative hazard rates can be written as

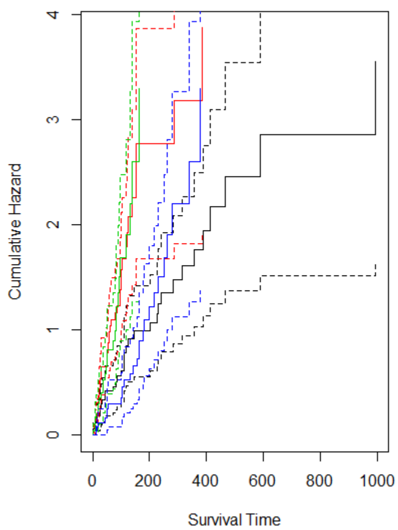
$$H(X) = \int_{u=0}^{x} h(u)du$$

If the hazard rates are proportional, which means the following h*(x) can be written as h*(x) = Ch(x), where h(x) is the previous rate and C is a constant number, then the associated cumulative hazard rate is

$$H^*(x) = \int_{u=0}^{x} h^*(u)du = \int_{u=0}^{x} Ch(u)du = C\int_{u=0}^{x} h(u)du = CH(x)$$

We proved that if the hazard rates are proportional, the cumulative hazard rates will also be proportional.

(b)



The cumulative hazard rate estimates assumption is reasonable.

(c)

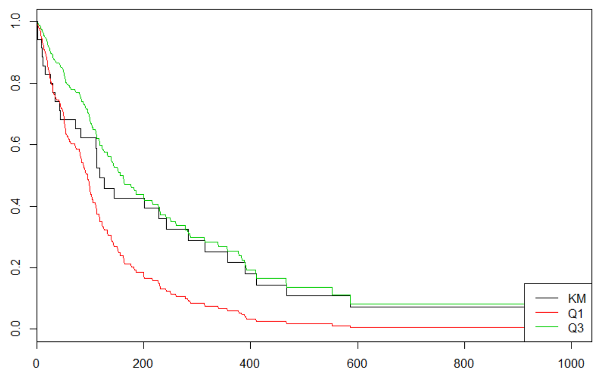
```
Analysis of Deviance Table
 Cox model: response is  Surv(stime)
 Model 1: ~ cell + Karn
 Model 2: ~ cell
   loglik  Chisq Df P(>|Chi|)
1 -512.23
2 -527.87 31.295  1 2.216e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Cox model: response is  Surv(stime)
 Model 1: ~ cell * Karn
 Model 2: ~ cell + Karn
   loglik Chisq Df P(>|Chi|)
1 -511.24
   2  -512.23 1.969  3    0.5789
```

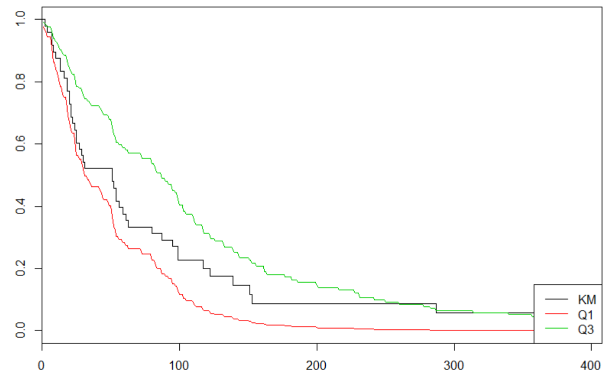
Based on the anova analysis results, we can conclude that model B is a significant improvement over model A, while model C is not a significant improvement over model B.

(d) For model B:

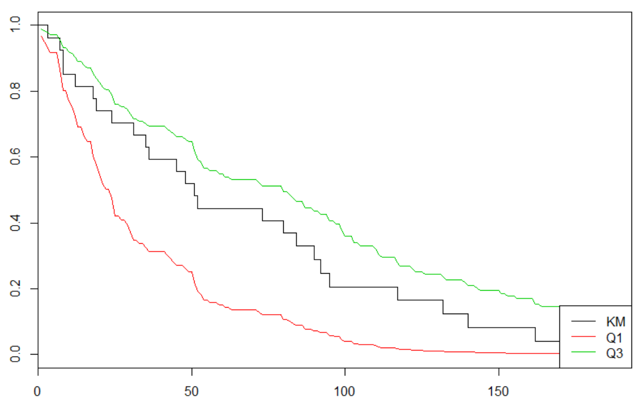
cell = 1



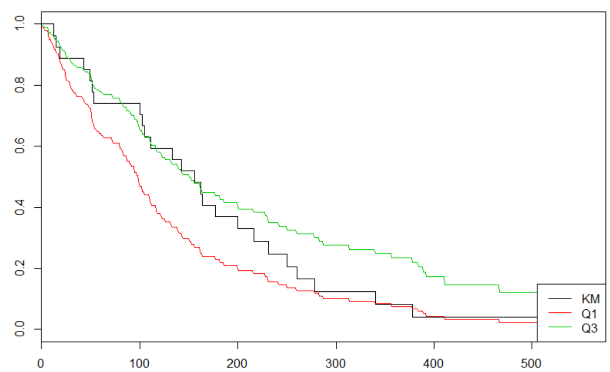
cell = 2



cell = 3

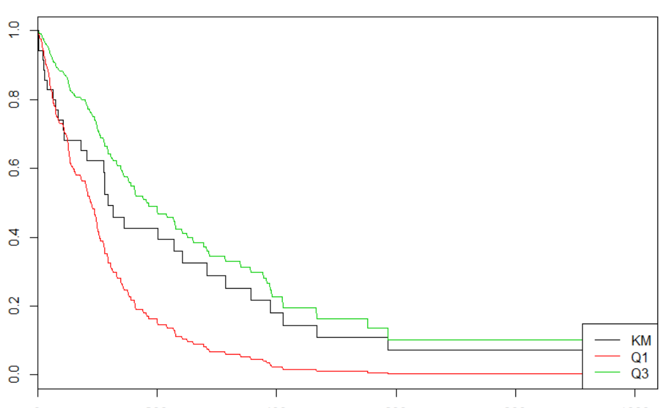


cell = 4

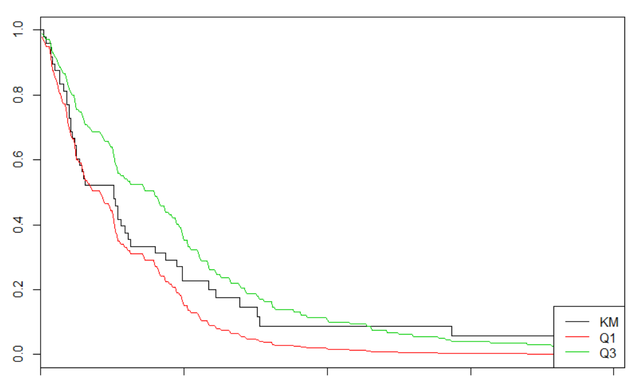


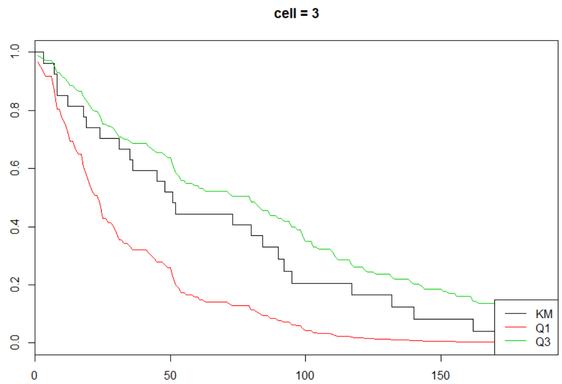
For model C:

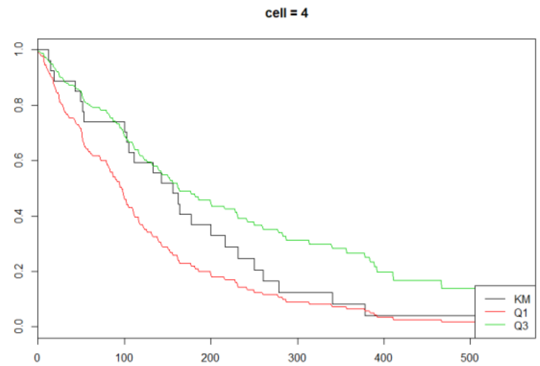
cell = 1



cell = 2







(e) Based on the result of (b) and (c), I don't think the relationship between a Karnofsky score and a patient's survival time differs by cell time. Thus, predictors **cell** and **Karn** are interactive in this way.

Q4

(a)

$$z_{1,}z_{2,}z_3 \text{ are conditional independent, so we have}$$

$$f\left(z_{1,}z_{2,}z_3\middle|\theta\right) = f(z_1|\theta)f(z_2|\theta)f(z_3|\theta)$$

Based on the information we have, k = 8.25. Thus

$$f\left(z_{1,}z_{2,}z_3\middle|\theta\right) = \frac{8.25}{\mu}\left(\frac{z_1}{\mu}\right)^{7.25} e^{-(\frac{z_1}{\mu})^{8.25}} \frac{8.25}{\mu}\left(\frac{z_2}{\mu}\right)^{7.25} e^{-(\frac{z_2}{\mu})^{8.25}} \frac{8.25}{\mu}\left(\frac{z_3}{\mu}\right)^{7.25} e^{-(\frac{z_3}{\mu})^{8.25}}$$

Since

$$\mu = \sqrt{(x-\theta_x)^2 + (y-\theta_y)^2}$$

We can write

$$f(z_1|\theta)$$

$$= \frac{8.25}{\sqrt{(-0.5-\theta_x)^2 + \left(0-\theta_y\right)^2}}\left(\frac{\frac{1}{0.926}}{\sqrt{(-0.5-\theta_x)^2 + \left(0-\theta_y\right)^2}}\right)^{7.25} e^{-\left(\frac{\frac{1}{0.926}}{\sqrt{(-0.5-\theta_x)^2+\left(0-\theta_y\right)^2}}\right)^{8.25}}$$

$$= \frac{e^{-\left(\frac{\frac{1}{0.926}}{\sqrt{(-0.5-\theta_x)^2+\left(0-\theta_y\right)^2}}\right)^{8.25}} 8.25/(\frac{1}{0.926})^{8.25}}{\left(\sqrt{(-0.5-\theta_x)^2+\left(0-\theta_y\right)^2}\right)^{8.25}}$$

$$f(z_2|\theta) = \frac{8.25}{\sqrt{(0.42-\theta_x)^2+\left(2-\theta_y\right)^2}}\left(\frac{\frac{1}{0.943}}{\sqrt{(0.42-\theta_x)^2+\left(2-\theta_y\right)^2}}\right)^{7.25} e^{-\left(\frac{\frac{1}{0.943}}{\sqrt{(0.42-\theta_x)^2+\left(2-\theta_y\right)^2}}\right)^{8.25}}$$

$$= \frac{e^{-\left(\frac{\frac{1}{0.943}}{\sqrt{(0.42-\theta_x)^2+\left(2-\theta_y\right)^2}}\right)^{8.25}} 8.25/(\frac{1}{0.943})^{8.25}}{\left(\sqrt{(0.42-\theta_x)^2+\left(2-\theta_y\right)^2}\right)^{8.25}}$$

$$f(z_3|\theta) \frac{8.25}{\sqrt{(1.5-\theta_x)^2+\left(1.27-\theta_y\right)^2}}\left(\frac{1/0.787}{\sqrt{(1.5-\theta_x)^2+\left(1.27-\theta_y\right)^2}}\right)^{7.25} e^{-(\frac{1/0.787}{\sqrt{(1.5-\theta_x)^2+\left(1.27-\theta_y\right)^2}})^{8.25}}$$

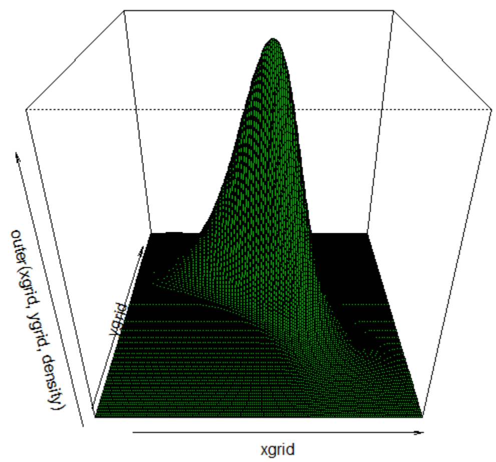
$$= \frac{e^{-\left(\frac{\frac{1}{0.787}}{\sqrt{(1.5-\theta_x)^2+\left(1.27-\theta_y\right)^2}}\right)^{8.25}} 8.25/(\frac{1}{0.787})^{8.25}}{\left(\sqrt{(1.5-\theta_x)^2+\left(1.27-\theta_y\right)^2}\right)^{8.25}}$$

$f\left(z_{1,}z_{2,}z_3\middle|\theta\right)$

$$= 8.25^3 \frac{e^{-\left(\frac{\frac{1}{0.926}}{\sqrt{(-0.5-\theta_x)^2+\left(0-\theta_y\right)^2}}\right)^{8.25}} (\frac{1}{0.926})^{8.25}}{\left(\sqrt{(-0.5-\theta_x)^2+\left(0-\theta_y\right)^2}\right)^{8.25}} \frac{e^{-\left(\frac{\frac{1}{0.943}}{\sqrt{(0.42-\theta_x)^2+\left(2-\theta_y\right)^2}}\right)^{8.25}} (\frac{1}{0.943})^{8.25}}{\left(\sqrt{(0.42-\theta_x)^2+\left(2-\theta_y\right)^2}\right)^{8.25}} \frac{e^{-\left(\frac{\frac{1}{0.787}}{\sqrt{(1.5-\theta_x)^2+\left(1.27-\theta_y\right)^2}}\right)^{8.25}} (\frac{1}{0.787})^{8.25}}{\left(\sqrt{(1.5-\theta_x)^2+\left(1.27-\theta_y\right)^2}\right)^{8.25}}$$

$$\pi\left(\theta\middle|z_{1,}z_{2,}z_3\right) = \frac{f\left(z_{1,}z_{2,}z_3\middle|\theta\right)\pi(\theta)}{\int f\left(z_{1,}z_{2,}z_3\middle|\theta\right)\pi(\theta)d\theta}$$

(b)



From (d) to (f), I think there might be something wrong with my code. I tried, but…

Q5

(a) We have

$$M_X(t) = E[e^{tX}]$$

We can write

$$\frac{dM_X(t)}{dt} = E[Xe^{tX}]$$

$$\frac{d^2M_X(t)}{d^2t} = E[X^2e^{tX}]$$

$$\dots$$

$$\frac{d^kM_X(t)}{d^kt} = E[X^ke^{tX}]$$

$$\left.\frac{d^kM_X(t)}{d^kt}\right|_{t=0} = E[X^ke^{tX}]\big|_{t=0} = E[X^k]$$

Proved.

(b)

$$m(t) = E[e^{tT(x)}] = \int e^{tT(x)}\, e^{\eta T(X) - A(\eta)} h(x)dx$$

$$m(t) = \int e^{(t+\eta)T(x) - A(\eta)}\, h(x)dx$$

Now

$$\int B(x|\eta)dx = 1$$

So

$$\int e^{(t+\eta)T(X) - A(\eta)}\, h(x)dx = 1$$

$$\int e^{\eta T(x)} h(x)dx = e^{A(\eta)}$$

$$m(t) = e^{-A(\eta)} \int e^{(\eta+t)T(x)} h(x)dx = e^{A(t+\eta) - A(\eta)}$$

$$\mathrm{m}'(\mathrm{t}) = e^{A(t+\eta) - A(\eta)} A'(t+\eta)$$

$$\mathrm{E(T)} = m'(t)|_{t=0} = A'(\eta)$$

$$\mathrm{m}''(\mathrm{t}) = e^{-A(\eta)} \left\{ A''(t+\eta) e^{A(t+\eta)} + e^{A(t+\eta)} [A'(t+\eta)]^2 \right\}$$

$$\mathrm{E(x^2)} = m''(t)|_{t=0} = A''(\eta) + [A'(\eta)]^2 = A''(\eta) + [E(x)]^2$$

$$\mathrm{v(x)} = A''(\eta)$$

Proved.

(c)

Log-likelihood

$$\mathrm{l}(\eta, \mathrm{x}) = log B(x|\eta)$$

We have

$$\frac{dl(\eta, x)}{d\eta} = \frac{d}{d\eta}\{\eta T(x) - A(\eta) + log h(x)\} = T(x) - A'(\eta)$$

$$\frac{d^2 l(\eta, x)}{d\eta^2} = -A''(\eta) = -v(x) < 0$$

Solution to

$$\frac{dl(\eta, x)}{d\eta} = 0$$

$\hat{\eta}$ satisfying

$$T(x) = A'(\eta) = E[T(x)]$$

$$uniquely\ maximizes\ \frac{d^2 l(\eta, x)}{d\eta^2} < 0$$

So $\hat{\eta}$ uniquely maximizes the log-likelihood.