CITY UNIVERSITY OF HONG KONG
STUDENTS' UNION
http://www.cityusu.net/

Name: CHAN King Yeung
SID: 1155119394

STAT 4006

Problem Sheet 3

Question 1

a) let $E(Y) = \alpha + \beta_1 x + \beta_2 x^2$, we maximise / minimise $E(Y)$ at $\frac{\partial E(Y)}{\partial x} = 0 \Rightarrow x = -\frac{\beta_1}{2\beta_2}$. If we drop x from $E(Y)$, $E(Y)$ will only be maximised / minimised at $x = 0$. Thus, models with lower-order term enable $E(Y)$ be maximised / minimised at $x \in \mathbb{R}$.

b) let $E(Y) = \alpha + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_1 x_2$, we have $E(Y) = \alpha + \beta_2 x_2$ when $x_1 = 0$. x_2 is still contributing to $E(Y)$ when $x_1 = 0$. If we drop x_2 from $E(Y)$, $E(Y)$ becomes α only when $x_1 = 0$. The effect of x_2 will be eliminated in the situation. Thus, models with lower-order term retain the effect of $x_i \ \forall i$.

Question 2

for $Y \sim \Gamma(k, \frac{\mu}{k})$, $f(y) = \frac{(k/\mu)^k}{\Gamma(k)} e^{-\frac{ky}{\mu}} y^{k-1}$

$$= \exp\left\{ \frac{y(-\frac{1}{\mu}) - \ln(\mu)}{\frac{1}{k}} + (k-1)\ln(y) + k\ln(k) + \ln[\Gamma(k)] \right\}$$

$\theta = -\frac{1}{\mu}$; $b(\theta) = -\ln(-\theta)$; $a(\phi) = \frac{1}{k}$; $c(y;\phi) = (\frac{1}{\phi} - 1)\ln(y) - \frac{\ln(\phi)}{\phi} - \ln[\Gamma(\frac{1}{\phi})]$

$\Rightarrow Y$ belongs to exponential dispersion with natural parameter $-\frac{1}{\mu}$

Question 3

We know fit.nb.colour2 is simpler than fit.nb.colour since it has 2 less parameters. For the likelihood ratio test comparing these 2 models, the deviance is 0.3834 and p-value is 0.8256. The test suggests fit.nb.colour2 fit better while holding fit.nb.colour.

$\beta = -0.2689$ denotes the effect of colour in fit.nb.colour2 which indicates the number of satellites decreases 0.2689 when the colour goes to next level or darker.

for the likelihood ratio test for $\beta = 0$, the deviance of fit.nb.colour2 and null model is 4.73 and p-value is 0.0296. The test suggests β is significant to explain the given data.

Question 4

a) Changing from 1 decade to the following decade will decrease the percentage of times a pitcher pitched a complete game by 6.94%.

b) $\hat{\pi}(13) = 0.7578 - 0.0694(13)$
$= -0.1444$

the predicted percentage of complete game for 2020-2029 is -14.44%.

c) $\hat{\pi}(13) = \frac{e^{1.148 - 0.315(13)}}{1 + e^{1.148 - 0.315(13)}}$
≈ 0.0499

logit link is more preferred since it restricts the probability to fall within $[0,1]$ which is more plausible.

Question 5

a) when $LI = 8$, $\text{logit}(\hat{\pi}) = -3.7771 + 0.1449(8)$
$\ln(\frac{\hat{\pi}}{1-\hat{\pi}}) = -2.6179$
$\hat{\pi} \approx 0.068$

b) when EL_{50}, $x = -\frac{-3.7771}{0.1449}$
≈ 26.06694

$\Rightarrow \hat{\pi} = 0.5$ when $x = 26.06694$

c) when $LI = 8$, rate of change $= 0.1449(0.068)(0.932)$
≈ 0.009

when $LI = 26.06694$, rate of change $= 0.1449(0.5)(0.5)$
≈ 0.036

d) when $LI = 14$, $\text{logit}(\hat{\pi}) = -3.7771 + 0.1449(14)$
$\ln(\frac{\hat{\pi}}{1-\hat{\pi}}) = -1.7485$
$\hat{\pi} \approx 0.15$

con't d) when $LI = 28$. $\text{logit}(\hat{\pi}_0) = -3.7771 + 0.1449(28)$

$$\ln\left(\frac{\hat{\pi}}{1-\hat{\pi}}\right) = 0.2801$$

$$\hat{\pi}_0 \approx 0.57$$

$\Rightarrow \hat{\pi}$ increases by 0.42

e) $$\frac{\hat{\pi}_0(x+1)}{1-\hat{\pi}_0(x+1)} = e^{-3.7771+0.1449(x+1)}$$

$$= e^{0.1449} \frac{\hat{\pi}_0(x)}{1-\hat{\pi}_0(x)}$$

$\Rightarrow$ the estimated odds of remission multiply by $e^{0.1449} \approx 1.156$ for an unit change in LI

f) $e^{0.1449 \pm 1.96(0.0593)}$

$\approx (1.0291, 1.2984)$

We infer that each unit increases in LI has at least 2.91% increase and at most 29.84% increase in the odds that a patient achieved remission.

g). $H_0: \beta = 0$ vs $H_1: \beta \neq 0$

$G^2 = 34.372 - 26.073$

$= 8.299$

critical value $= 3.841$

Since $G^2 > 3.841$, we reject H_0 at $\alpha = 0.05$.

Question 6

a) for $Y \sim \text{Bern}(p)$,

$$L(p) = p^{\sum y_i}(1-p)^{n-\sum y_i}$$

$L(p)$ also is a kernel function.

for $Y \sim b(n,p)$,

$$L(p) = \prod \binom{n_i}{y_i} p^{\sum y_i}(1-p)^{n-\sum y_i}$$

kernel function $= p^{\sum y_i}(1-p)^{n-\sum y_i}$ which is the same as bernoulli

b) the number of parameters in saturated model equals the number of observations. For n bernoulli data, it has n parameters $\{\pi_{ij}\}$; for N binomal data, it has N parameter $\{\pi_1, \dots, \pi_N\}$. Thus, their kernel functions are different. The deviance contains the saturated models also differ.

c) the difference of the deviance cancel out the common saturated model and only rely on kernel functions of unsaturated models. where a) shows that the kernel functions of unsaturated independent of the data entry.