BM_hw1

PROBALEM 1

```
does = c(0, 1, 2, 3, 4)
number = c(30, 30, 30, 30, 30)
killed = c(2, 8, 15, 23, 27)
data1 = data.frame(does, number, killed)

#data preparation
x = does
y =killed
n = number
resp = cbind(y, n-y) # dead = 0, survived = 1
```

fit models

```
glm_logit=glm(resp~x, family=binomial(link='logit'))
glm_probit=glm(resp~x, family=binomial(link='probit'))
glm_clog=glm(resp~x, family=binomial(link='cloglog'))
```

#1.1

```
# estimate of beta
#1
summary(glm_logit)
beta1 = glm_logit$coefficients[2]
se1=sqrt(vcov(glm_logit)[2,2])
beta1+c(qnorm(0.025),-qnorm(0.025))*se1
sum(residuals(glm_logit,type='deviance')^2) # deviance
predict(glm_logit, data.frame(x=0.01),type='response')

#2
summary(glm_probit)
beta2= glm_probit$coefficients[2]
se2 = sqrt(vcov(glm_probit)[2,2])
beta2+c(qnorm(0.025),-qnorm(0.025))*se2
sum(residuals(glm_probit,type='deviance')^2) # deviance
predict(glm_probit, data.frame(x=0.01),type='response')


#3
summary(glm_clog)
beta3= glm_clog$coefficients[2]
se3 = sqrt(vcov(glm_clog)[2,2])
beta3+c(qnorm(0.025),0,-qnorm(0.025))*se3
sum(residuals(glm_clog,type='deviance')^2)# deviance
predict(glm_clog, data.frame(x=0.01),type='response')
```

Model	beta est	ci beta	deviance	p(dying
logit	1.16	(0.806,1.517)	0.378	0.0901
probit	0.6863805	(0.497, 0.876)	0.314	0.0853
c-log-log	0.7468193	(0.532,0.961)	2.23	0.128

comments:based on the results, deviance of model with probit link is smallest, so it is optimal model.

#1.2

to calculate LD50, $\beta_0 + \beta_1 x_0 = g(0.5)$ point estimates for model with logit or probit link are same: $x_0 = -\frac{\beta_0}{\beta_1}$ $\frac{\partial x_0}{\partial \beta_0} = -\frac{1}{\beta_1}$ $\frac{\partial x_0}{\partial \beta_1} = \frac{\beta_0}{\beta_1^2}$

c-log-log point estimate: $log(-log(1-0.5)) = \beta_0 + \beta_1 x_0$ $x_0 = \frac{log(-log(1-0.5))-\beta_0}{\beta_1}$ $\frac{\partial x_0}{\partial \beta_0} = -\frac{1}{\beta_1}$ $\frac{\partial x_0}{\partial \beta_1} \frac{\beta_0 - log(log2)}{\beta_1^2}$

asymptotic variance of $\bar{x}_0$: $var(\hat{x}_0) = (\frac{\partial x_0}{\partial \beta_0})^2 var(\hat{\beta}_0) + (\frac{\partial x_0}{\partial \beta_1})^2 var(\hat{\beta}_1) + 2\frac{\partial x_0}{\partial \beta_0})(\frac{\partial x_0}{\partial \beta_1})cov(\hat{\beta}_0, \hat{\beta}_1)$

From r output, point estimate of LD50 for model with logit and probit link is 7.389 and 8.841 for model with c-log link.
90% CI LD50 for model with logit and probit link is (5.509631, 9.909583)
90% CI LD50 for model with c-log-log link is (6.526261, 11.977407)

PROBLEM2

```
amount = c(10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90)
offers = c(4, 6, 10, 12, 39, 36, 22, 14, 10, 12, 8, 9, 3, 1, 5, 2, 1)
enrolls = c(0, 2, 4, 2, 12, 14, 10, 7, 5, 5, 3, 5, 2, 0, 4, 2, 1)
df1 = c(amount, offers, enrolls)

x1 = amount
y1 = enrolls
n1 = offers
resp1 = cbind(y1, n1-y1)
```

#2.1

```
glm_logit1 = glm(resp1 ~ x1, family = binomial(link = 'logit'))
dev = sum(residuals(glm_logit1,type='deviance')^2)

1-pchisq(dev, 15) # p value = 1- 0.2204655, cannot reject the null hypothesis
```

Deviance of the logistic model is 10.6 which follows the chi-square(15). The p value is 0.779 which is greater than the significant levels. Therefore, we cannot reject the null hypothesis that the proposed model is true model.

###2.2

```
summary(glm_logit1)
predict(glm_logit1, se.fit=TRUE)
beta_1 = glm_logit1$coefficients[2]
se1 = sqrt(vcov(glm_logit1)[2,2])
beta_1+c(qnorm(0.025),-qnorm(0.025))*se1
```

The log odds of enrolls increase 0.03 with one thousand scholarship increases.

The log odds of enrolls is -1.647 when scholarship is 0.

We are 95% confident the coefficients of beta_1 is between 0.01197845 and 0.04992240.

2.3

```
beta_0 = glm_logit1$coefficients[1]
betacov1=vcov(glm_logit1)
x_fit = (log(2/3)-beta_0)/beta_1
var_x=betacov1[1,1]/(beta_1^2)+(log(2/3)-beta_0)^2*betacov1[2,2]/beta_1^4+2*(log(2/3)-beta_0)*betacov1[

c(x_fit,sqrt(var_x))
x_fit+c(qnorm(0.025),-qnorm(0.025))*sqrt(var_x)
```

$\beta_0 + \beta_1 x_0 = g(0.4)$

To get 40% yield rate, we should provide \$40,134 scholarship. We are 95% confident that we should provide between \$30,583 to \$49,685 amounts of scholarship to get 40% yield rate.