Stat 516 Homework 5 Solutions

1. Consider conducting m hypothesis tests. Let V denote the number of type I errors. Let R denote the number of rejected null hypotheses. Let $Q = V/R$ if $R > 0$, and let $Q = 0$ if $R = 0$. Let $\alpha \in (0, 1)$ be fixed. By definition, a method that strongly controls FDR at level α has the property that $E(Q) \leq \alpha$ no matter which or how many of the m null hypotheses are true or false. Note that if all null hypotheses are true, $Q = 1$ if $R > 0$, $Q = 0$ if $R = 0$, and $V = R$. Thus,

$$\begin{aligned} E(Q) &= E(Q|R > 0)P(R > 0) + E(Q|R = 0)P(R = 0) \\ &= 1 * P(R > 0) + 0 * P(R = 0) = P(R > 0) = P(V > 0) = \text{FWER}, \end{aligned}$$

when all null hypotheses are true. Thus, when all null hypotheses are true, FDR control ($E(Q) \leq \alpha$) implies FWER control (FWER$\leq \alpha$).

2. (a) By definition of conditional probability,

$$P(H_2|H_1) = \frac{P(H_1H_2)}{P(H_1)}.$$

Now note that

$$\begin{aligned} P(H_1H_2) &= P(H_1H_2|A_1A_2)P(A_1A_2) + P(H_1H_2|A_1B_2)P(A_1B_2) \\ &\quad + P(H_1H_2|B_1A_2)P(B_1A_2) + P(H_1H_2|B_1B_2)P(B_1B_2) \\ &= 0.5^2 * 0.7 * 0.9 + 0.5 * 1 * 0.7 * 0.1 + 1 * .5 * 0.3 * 0.6 + 1 * 1 * 0.3 * 0.4 \\ &= 0.4025 \end{aligned}$$

and

$$\begin{aligned} P(H_1) &= P(H_1|A_1)P(A_1) + P(H_1|B_1)P(B_1) \\ &= 0.5 * 0.7 + 1 * 0.3 = 0.65. \end{aligned}$$

Thus, $P(H_2|H_1) = 0.4025/0.65 = 0.6192308$.

(b) It is most likely that coin A was used for both flips. See the relevant calculations below.

$$P(A_1A_2|H_1H_2) = P(H_1H_2|A_1A_2)P(A_1A_2)/0.4025 = 0.5^2 * 0.7 * 0.9/0.4025 = 0.3913043$$

$$P(B_1A_2|H_1H_2) = P(H_1H_2|B_1A_2)P(B_1A_2)/0.4025 = 1 * .5 * 0.3 * 0.6/0.4025 = 0.2236025$$

$$P(A_1B_2|H_1H_2) = P(H_1H_2|A_1B_2)P(A_1B_2)/0.4025 = 0.5 * 1 * 0.7 * 0.1/0.4025 = 0.08695652$$

$$P(B_1B_2|H_1H_2) = P(H_1H_2|B_1B_2)P(B_1B_2)/0.4025 = 1 * 1 * 0.3 * 0.4/0.4025 = 0.2981366$$

3.
```
#Density of the absolute value of a noncentral
#t random variable with df degrees of freedom
#and noncentrality parameter ncp.

f.abs.t=function(t,df,ncp=0){
```

```
  (t>0)*(dt(t,df=df,ncp=ncp)+dt(-t,df=df,ncp=ncp))
}

#Density of the p-value from a t-test
#when the noncentrality parameter is ncp
#and the degrees of freedom are df.

f=function(p,df,ncp=0)
{
  tt=qt(1-p/2,df=df)
  f.abs.t(tt,df=df,ncp=ncp)/f.abs.t(tt,df=df)
}

f(.1,18,1.6)
[1] 2.072947
```

4. (a) $N_R = 2.12 + 2.23 + 4.42 + 2.14 = 10.91 \quad N - N_H = 10 - 4 = 6$

i	Ordered Stat	$P_{\text{hit}}(S,i)$	$P_{\text{miss}}(S,i)$	$P_{\text{hit}}(S,i) - P_{\text{miss}}(S,i)$
1	-4.42	4.42/10.91	0	4.42/10.91
2	-2.14	6.56/10.91	0	6.56/10.91
3	-2.12	8.68/10.91	0	8.68/10.91
4	-0.25	8.68/10.91	1/6	8.68/10.91-1/6
5	-0.17	8.68/10.91	1/3	8.68/10.91-1/3
6	-0.14	8.68/10.91	1/2	8.68/10.91-1/2
7	0.26	8.68/10.91	2/3	8.68/10.91-2/3
8	0.36	8.68/10.91	5/6	8.68/10.91-5/6
9	0.60	8.68/10.91	1	8.68/10.91-1
10	2.23	1	1	0

```
           Phit     Pmiss  Phit-Pmiss
 [1,] 0.4051329 0.0000000  0.40513291
 [2,] 0.6012832 0.0000000  0.60128323
 [3,] 0.7956004 0.0000000  0.79560037
 [4,] 0.7956004 0.1666667  0.62893370
 [5,] 0.7956004 0.3333333  0.46226703
 [6,] 0.7956004 0.5000000  0.29560037
 [7,] 0.7956004 0.6666667  0.12893370
 [8,] 0.7956004 0.8333333 -0.03773297
 [9,] 0.7956004 1.0000000 -0.20439963
[10,] 1.0000000 1.0000000  0.00000000
```

(b) $M(S) = 0.7956, m(S) = -0.2044$

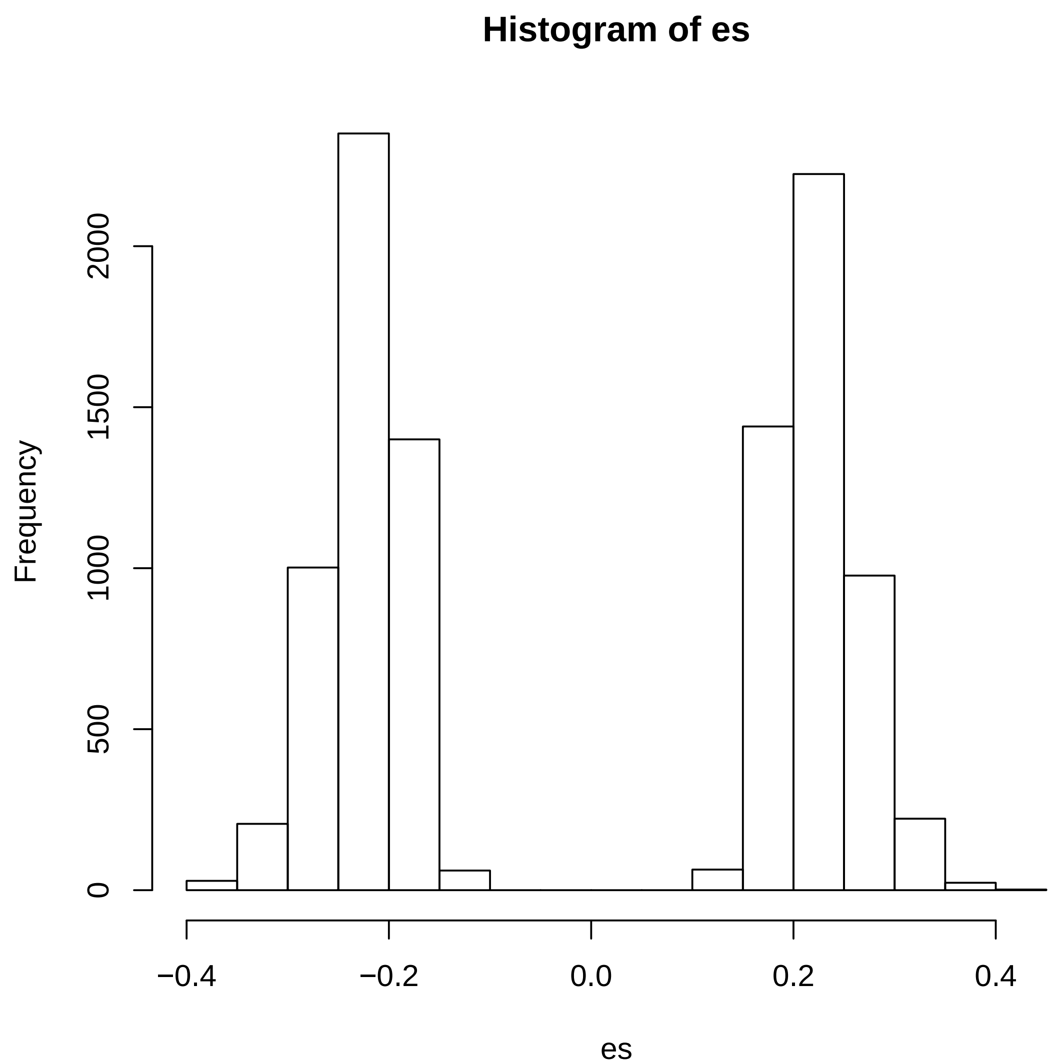
(c) $ES(S) = 0.7956$

(d)
```
getES=function(r,S,p=1,plt=F)
{
  N=length(r)
  NH=length(S)
  inS=(1:N)%in%S
```

```
  Pmiss=(!inS)/(N-NH)
  Phit=inS*abs(r)^p/sum(abs(r[inS])^p)
  ordr=order(r)
  Pmiss=cumsum(Pmiss[ordr])
  Phit=cumsum(Phit[ordr])
  Phit.Pmiss=Phit-Pmiss
  mM=range(Phit.Pmiss)
  Mgenm=(mM[2]>=-mM[1])
  ES=mM[2]*Mgenm+mM[1]*(1-Mgenm)
  if(plt){
    plot(1:N,Phit.Pmiss)
    lines(1:N,Phit.Pmiss)
  }
  ES
}
```

(e)

```
n=10000
es=rep(0,n)
for(i in 1:n){
  es[i]=getES(rnorm(1000),1:100)
}
hist(es)
```



(f) No simulation is actually necessary here. It is not difficult to see that $M(S)$ and $m(S)$ are guaranteed to be 0.5 and -0.5, respectively.

(g) Yes. An enrichment score of 0.5 or -0.5 is more extreme (farther from 0) than the statistics in part (e).