Experiments With Periodicity Tests

Summary

We compare Monte-Carlo tests using a generalization of the g-statistic and the likelihood ratio test. From Walker (2003), we know that these tests should be asymptotically equivalent. The harmonic regression, see below, is simulated with various snr and series lengths n. The number of frequencies used, nf = 50, when $n = 11$, 25; nf = 100 for $n = 75$ and nf = 200 for $n = 100$. For each parameter setting NSIM = 1000 simulations were done.

A paired two-sample t-test was used to test the equality of the p-values for each parameter setting. No differences were found. The periodogram approach is simpler, so it will be used in future work.

The power of a 5% test is tabulated in each case as well. The two Monte-Carlo tests are almost identical and clearly outperform the Fisher exact test.

The boxplots also illustrate that the performance of the MC tests is about the same and better than the Fisher exact test. When snr=0, the p-values in all tests are uniformly distributed on (0, 1) as expected.

Another fact that emerges from the simulations is that even for $n = 100$, Fisher's exact g test may not work very well.

Background

Standard Definition of the Periodogram

The periodogram provides an unbiased estimate of the spectral density function. The periodogram is usually computed at the Fourier frequencies, $f_j = j/n$ where $j = [-(n-1)/2], \ldots, 0, 1, \ldots, [n/2]$.

$$I(f_j) = \frac{1}{n}\left(\left|\sum_{t=1}^{n} z_t\, e^{2\pi i f_j (t-1)}\right|\right)^2$$

The periodogram provides an unbiased estimate of the spectral density function. Another interpretation of $I(f_j)$ is that it is proportional to the coefficient of determination of the regression of the data z_t on a sinusoid with frequency f_j. The sinusoids at frequencies f_{j_1} and f_{j_2} are uncorrelated so the coefficient of determination does not depend on which other sinusoids are the Fourier frequencies are included.

Fisher's g Test

Fisher's g statistic can be written,

$$g = \frac{\max_j I(f_j)}{\sum_{j=1}^{m} I(f_j)}$$

where m is $(n-1)/2$ or $(n-2)/2$ according as n is odd or even.

Then under the assumption that e_t is $\text{NID}(0, \sigma^2)$ and the null hypothesis $\mathcal{H}_0 : \lambda = 0$ the cdf of g is given by,

$$F(x) = \Pr\{g \leq x\} = 1 - \sum_{i=1}^{\rho} \binom{n}{i} (-1)^i (1 - i\,x)^{n-i}$$

where $\rho = \lfloor 1/x \rfloor$. An upper tail test is used so the observed p-value is given by $1 - F(g)$.

Note that although the "official" null hypothesis is $\mathcal{H}_0 : \lambda = 0$, Fisher's g test will also detect any any departure from an uncorrelated sequence provided the sample is large enough. The test will be optimal against an alternative hypothesis $\mathcal{H}_1 : \lambda = \lambda_j$ where $\lambda_j \in \{1/n, 2/n, \ldots, [n/2]/n\}$. When n is small this restriction to the Fourier frequencies is unrealistic.

Monte-Carlo Tests

LR Method

The simple harmonic regression model may be written,

$$z_t = \mu + A\cos(2\pi\lambda t) + B\sin(2\pi\lambda t) + e_t, \;\; e_t \sim \text{IID}(0, \sigma^2)$$

where A and B jointly determine the amplitude and phase of the sinusoid and λ is the frequency. The parameters are estimated by MLE and then the loglikelihood ratio statistic, LR, is evaluated.

The signal-to-noise ratio is defined by

$$\text{snr} = (A^2 + B^2)/\sigma^2$$

The larger snr, the easier it is to detect the signal.

Fisher g statistic

A generalized version of the periodogram may be defined by,

$$I(f) = \frac{1}{n_T}\left(\left|\sum_{t \in T} z_t\, e^{2\pi i f_j (t-1)}\right|\right)^2$$

where $0 < f < 1/2$, T is the set of time points at which the series is observed and n_T is the number of time points. Then

$$g = \max I(f)/S$$

where S is the total corrected sum of squares,

$$S = \sum_{t \in T} (z_t - \bar{z})^2$$

and $\bar{z}$ is the sample average.

Difference Between Periodogram and Harmonic Regression Tests

Purpose

As noted in Walker (2003) and other references, the periodogram and harmonic regression estimates for frequency are asymptotically equivalent. We use simulation to investigate if there is any difference in small samples.

The standard Fisher g test is also compared to the two Monte-Carlo tests.

Experiment 1. n = 11, nf = 50

n is the length of series
nf is the number of frequencies used

R Script

```
#Comparing Fisher's g statistic with LLR and exact Fisher test
start.time<-proc.time()[1]
n<-11
f<-1.5/n
Numf<-50
NSIM<-1000
snrS<-c(0,5,10,15,20,25)
tb<-array(numeric(length(snrS)*NSIM*3),dim=c(length(snrS),3,NSIM))
dimnames(tb)<-list(snrS, c("g", "LR","F"), 1:NSIM)
for (isnr in 1:length(snrS)) {
    snr<-snrS[isnr]
    A<-sqrt(snr)/3
    B<-sqrt(snr-A^2)
    pF<-pg<-pLR<-numeric(NSIM)
    for (isim in 1:NSIM){
        z<-SimulateHReg(n=n,f=f,A=A,B=B,phi=0, simPlot=FALSE)
        pg[isim]<-PeriodicityTest(z, Numf=Numf, which.test="g", nullDis="b")
        pLR[isim]<-PeriodicityTest(z, Numf=Numf, which.test="LLR", nullDis="b")
        pF[isim]<-FisherGTest(z)$p.value
    }
    tb[isnr,1,]<-pg
    tb[isnr,2,]<-pLR
    tb[isnr,3,]<-pF
}
end.time<-proc.time()[1]
total.time<-end.time-start.time
#
#plot results and tabulate t-tests
fsnr<-ordered(snrS)
fTest<-ordered(c("g","LR","F"))
nsnr<-length(snrS)
nTest<-length(fTest)
tb.df<-data.frame(pvalue=c(tb),SNR=rep(fsnr, NSIM*nTest),
           TEST=rep(rep(fTest, rep(rep(nsnr,nTest))),NSIM))
ti<-paste("Comparison for various snr's with n =",n)
bwplot(TEST~pvalue | SNR, data=tb.df,main=ti)
bwplot(TEST~pvalue | SNR, data=tb.df, subset=TEST!="F",
           scales=list(x=list(relation="free")),main=ti)
ctb<-matrix(numeric(3*nsnr),ncol=3)
for (i in 1:nsnr) {
    out<-t.test(pvalue~TEST, data=tb.df, subset=TEST!="F"&SNR==fsnr[i],paired=TRUE)
    ctb[i,]<-c(out$conf.int,out$p.value)
    }
dimnames(ctb)<-list(snrS, c("lower","upper","pvalue"))
```

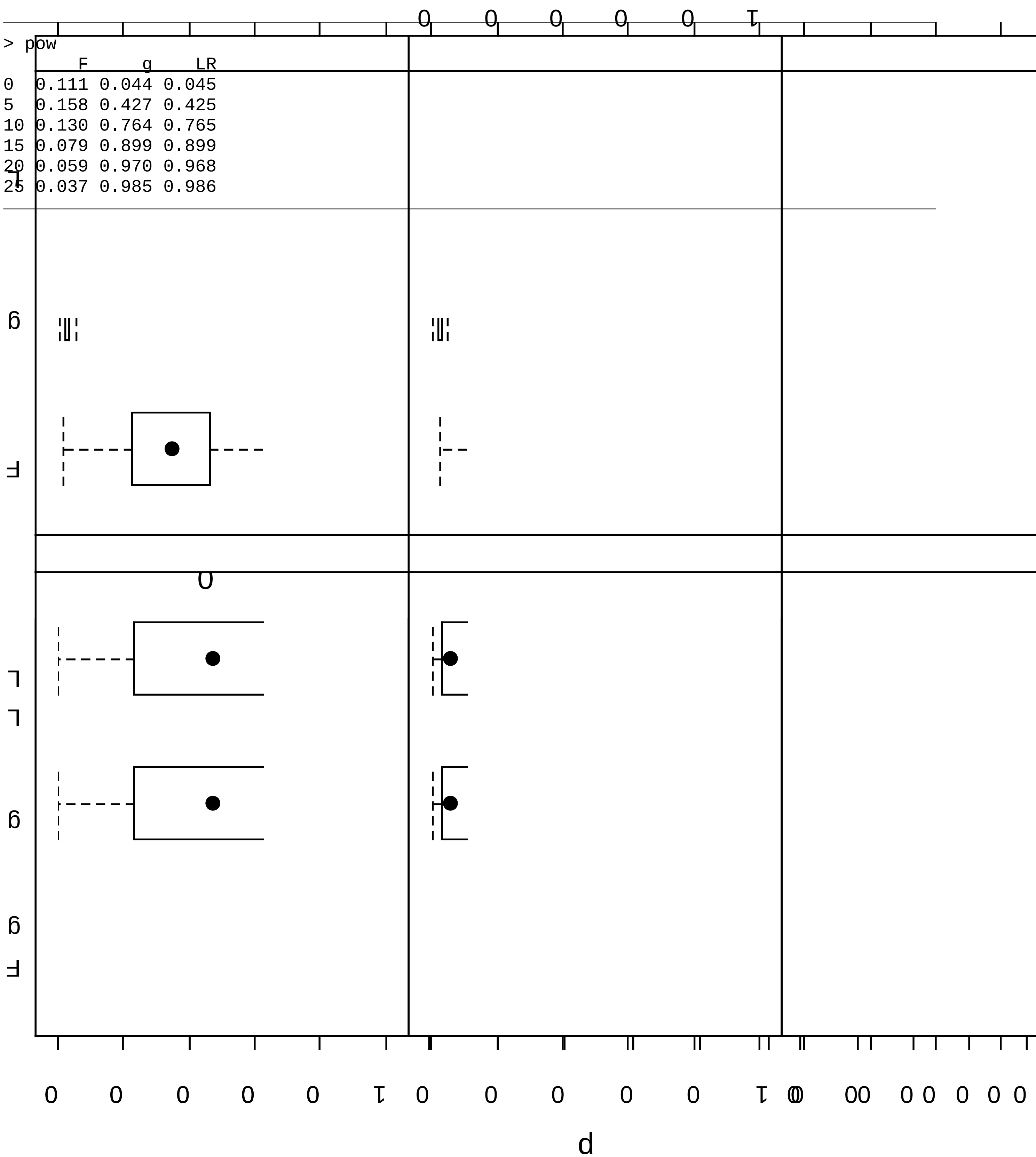
Results

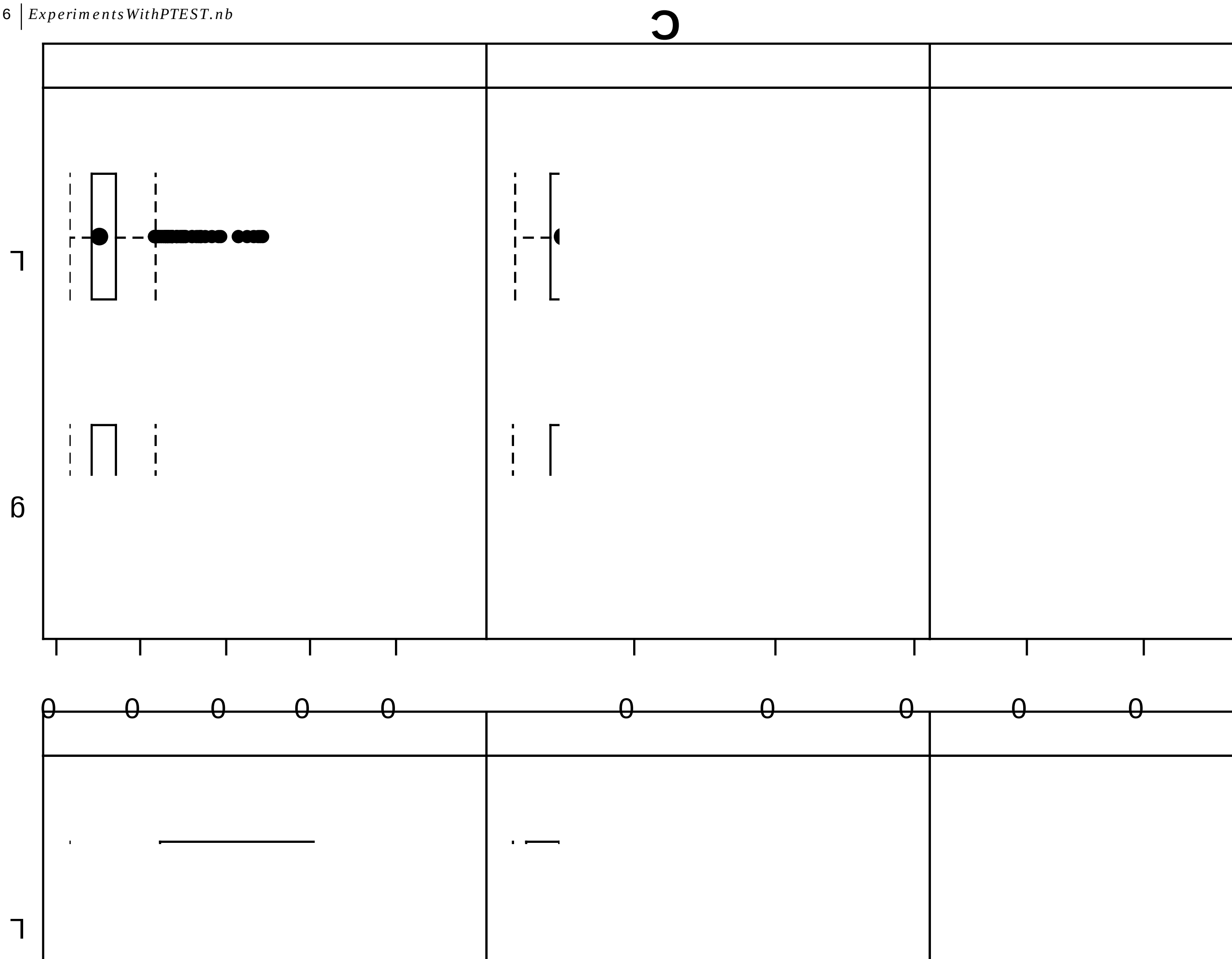
F - for Fisher's g test (exact)

g - for generalized periodogram statistic (Monte Carlo test)

LR - for likelihood-ratio test statistic (Monte Carlo test)

```
> pow
      F     g     LR
0  0.111 0.044 0.045
5  0.158 0.427 0.425
10 0.130 0.764 0.765
15 0.079 0.899 0.899
20 0.059 0.970 0.968
25 0.037 0.985 0.986
```





- **Paired t-test for difference in power**

```
> ctb
          lower        upper    pvalue
0  -1.535638e-04 0.0005829208 0.2528939
5  -3.834341e-04 0.0001054619 0.2648057
10 -9.285573e-05 0.0002730377 0.3341064
15 -1.880999e-04 0.0001015085 0.5575180
20 -1.169337e-04 0.0001585295 0.7670474
25 -1.156947e-04 0.0001534909 0.7829650
```

- **Conclusions**

Fisher's exact test doesn't work very well in this example. There is no difference in performance in the LR and g Monte-Carlo tests. Next we will examine for larger sample sizes.

■ Experiment 2. $n = 25$, nf = 50

■ S Script

```
#Comparing Fisher's g statistic with LLR and exact Fisher test
start.time<-proc.time()[1]
library(lattice)
n<-25
f<-1.5/n
Numf<-50
NSIM<-1000
snrS<-c(0,5,10,15,20,25)*(11/n)^2
tb<-array(numeric(length(snrS)*NSIM*3),dim=c(length(snrS),3,NSIM))
dimnames(tb)<-list(snrS, c("g", "LR","F"), 1:NSIM)
for (isnr in 1:length(snrS)) {
    snr<-snrS[isnr]
    A<-sqrt(snr)/3
    B<-sqrt(snr-A^2)
    pF<-pg<-pLR<-numeric(NSIM)
    for (isim in 1:NSIM){
        z<-SimulateHReg(n=n,f=f,A=A,B=B,phi=0, simPlot=FALSE)
        pg[isim]<-PeriodicityTest(z, Numf=Numf, which.test="g", nullDis="b")
        pLR[isim]<-PeriodicityTest(z, Numf=Numf, which.test="LLR", nullDis="b")
        pF[isim]<-FisherGTest(z,UseC=TRUE)$p.value
    }
    tb[isnr,1,]<-pg
    tb[isnr,2,]<-pLR
    tb[isnr,3,]<-pF
}
end.time<-proc.time()[1]
total.time<-end.time-start.time
#
#plot results and tabulate t-tests
fsnr<-ordered(snrS)
fTest<-ordered(c("g","LR","F"))
nsnr<-length(snrS)
nTest<-length(fTest)
tb.df<-data.frame(pvalue=c(tb),SNR=rep(fsnr, NSIM*nTest),
           TEST=rep(rep(fTest, rep(rep(nsnr,nTest))),NSIM))
write.table(tb.df, file="d:/r/2008/ptest/simg/tbn25Numf50.dat")
ti<-paste("Comparison for various snr's with n =",n)
bwplot(TEST~pvalue | SNR, data=tb.df,main=ti)
bwplot(TEST~pvalue | SNR, data=tb.df, subset=TEST!="F",
           scales=list(x=list(relation="free")),main=ti)
ctb<-matrix(numeric(3*nsnr),ncol=3)
for (i in 1:nsnr) {
    out<-t.test(pvalue~TEST, data=tb.df, subset=TEST!="F"&SNR==fsnr[i],paired=TRUE)
    ctb[i,]<-c(out$conf.int,out$p.value)
    }
dimnames(ctb)<-list(snrS, c("lower","upper","pvalue"))
attach(tb.df)
pow<-tapply(pvalue, INDEX=list(SNR, TEST), FUN=function(x) sum(x<0.05)/NSIM)
detach("tb.df")
savews()
```

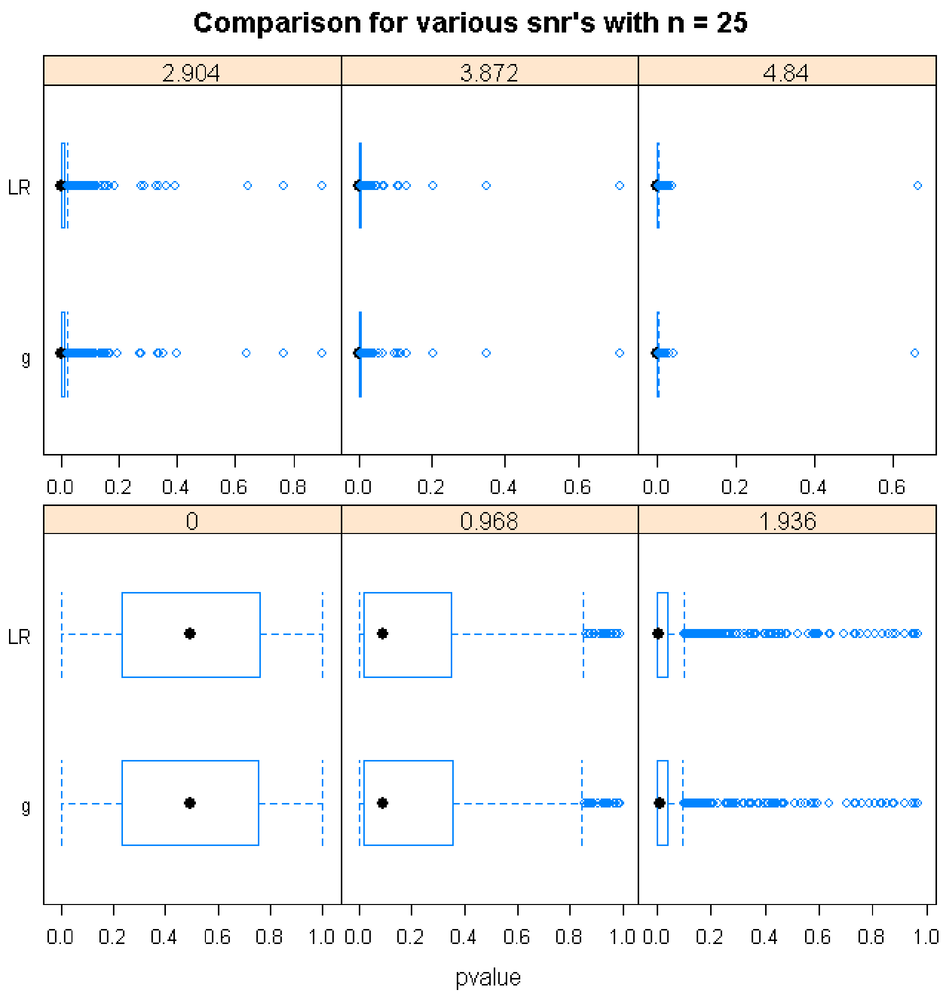
Power Comparison for 5% test

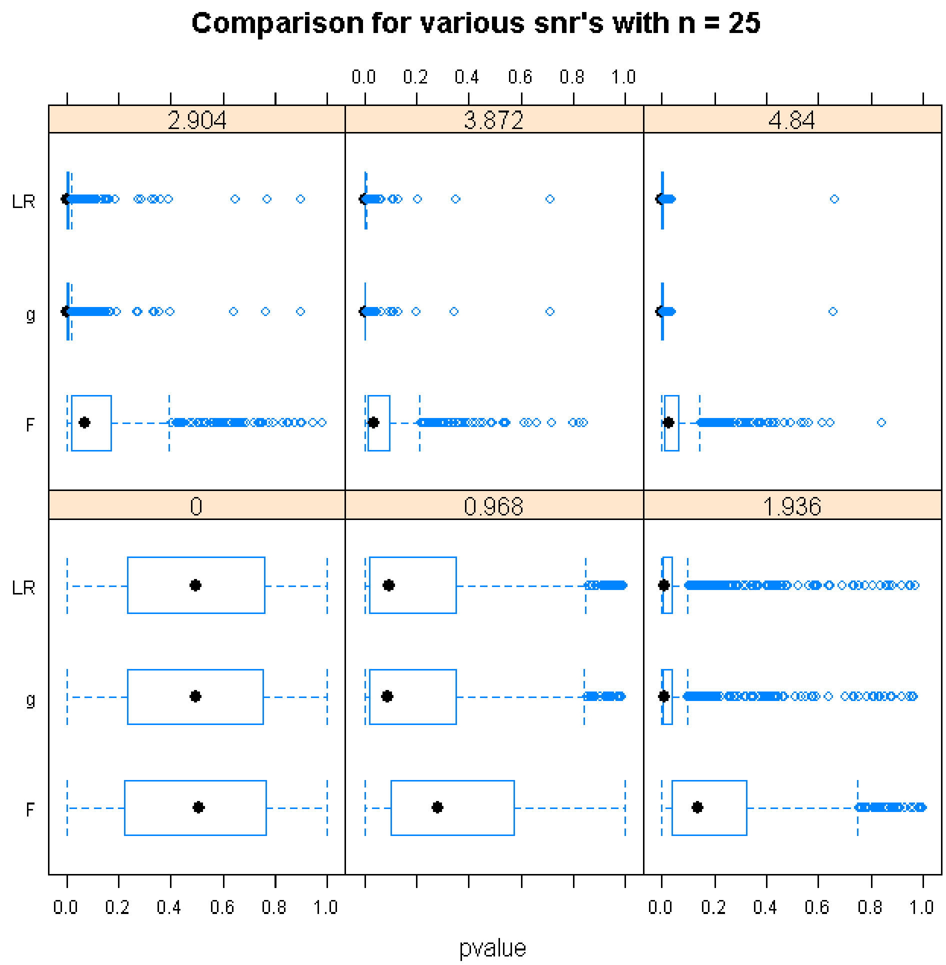
```
> pow
          F     g    LR
0     0.058 0.053 0.054
0.968 0.146 0.382 0.384
1.936 0.292 0.779 0.781
2.904 0.423 0.945 0.945
3.872 0.576 0.990 0.991
4.84  0.689 0.999 0.999
```

- **Paired t-test for difference in power**

```
> ctb
              lower        upper    pvalue
0     -3.051641e-04 4.199527e-04 0.7561349
0.968 -1.425606e-04 4.375311e-04 0.3186035
1.936 -1.511055e-04 2.104995e-04 0.7472800
2.904 -1.685636e-04 2.197821e-05 0.1314501
3.872 -4.754556e-05 4.934538e-05 0.9709293
4.84  -5.548136e-05 1.668524e-05 0.2917091
```

■ Boxplots


Comparison for various snr's with n = 25
2.904
3.872
4.84
LR
g
0.0 0.2 0.4 0.6 0.8
0.0 0.2 0.4 0.6
0.0 0.2 0.4 0.6
0
0.968
1.936
0.0 0.2 0.4 0.6 0.8 1.0
pvalue



■ Experiment 3. $n = 75$, nf = 100

```
#Comparing Fisher's g statistic with LLR and exact Fisher test
start.time<-proc.time()[1]
library(lattice)
n<-75
f<-1.5/n
Numf<-100
NSIM<-1000
snrS<-round(c(0,5,10,15,20,25)*(11/n)^2, 4)
tb<-array(numeric(length(snrS)*NSIM*3),dim=c(length(snrS),3,NSIM))
dimnames(tb)<-list(snrS, c("g", "LR","F"), 1:NSIM)
for (isnr in 1:length(snrS)) {
    snr<-snrS[isnr]
    A<-sqrt(snr)/3
    B<-sqrt(snr-A^2)
    pF<-pg<-pLR<-numeric(NSIM)
    for (isim in 1:NSIM){
        z<-SimulateHReg(n=n,f=f,A=A,B=B,phi=0, simPlot=FALSE)
        pg[isim]<-PeriodicityTest(z, Numf=Numf, which.test="g", nullDis="b")
        pLR[isim]<-PeriodicityTest(z, Numf=Numf, which.test="LLR", nullDis="b")
```

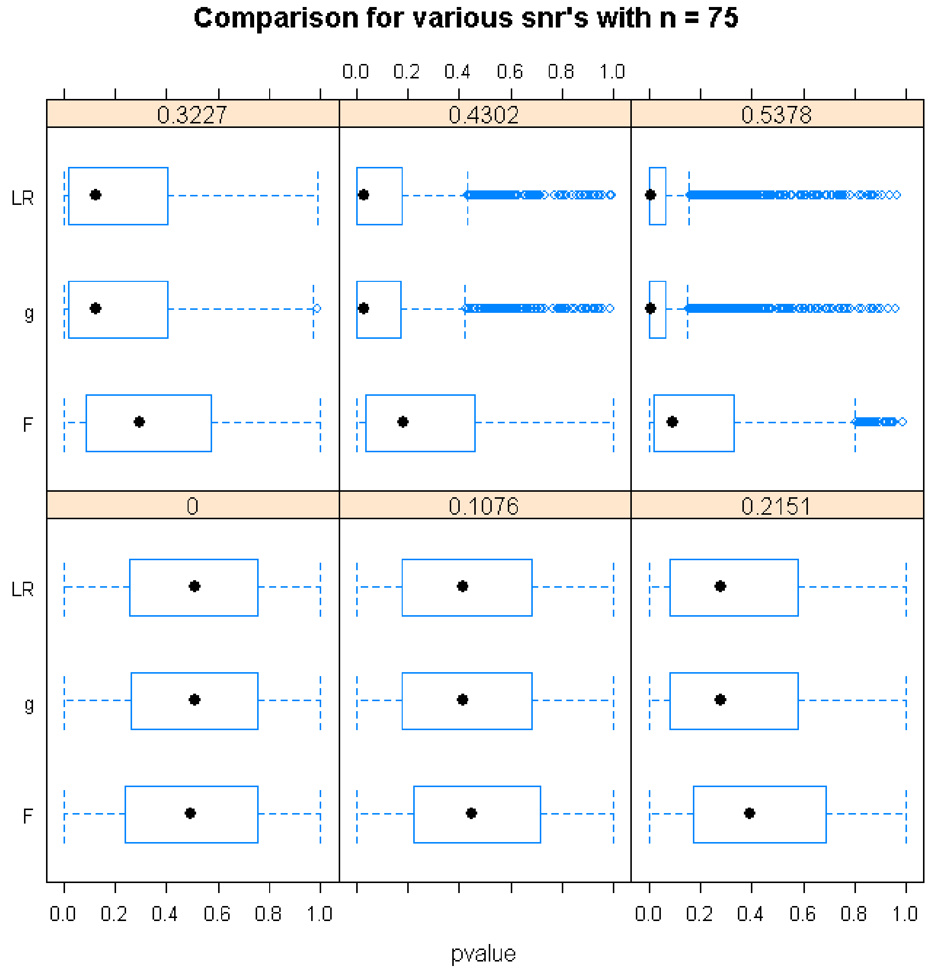
```
        pF[isim]<-FisherGTest(z,UseC=TRUE)$p.value
    }
    tb[isnr,1,]<-pg
    tb[isnr,2,]<-pLR
    tb[isnr,3,]<-pF
}
end.time<-proc.time()[1]
total.time<-end.time-start.time
#
#plot results and tabulate t-tests
fsnr<-ordered(snrS)
fTest<-ordered(c("g","LR","F"))
nsnr<-length(snrS)
nTest<-length(fTest)
tb.df<-data.frame(pvalue=c(tb),SNR=rep(fsnr, NSIM*nTest),
           TEST=rep(rep(fTest, rep(rep(nsnr,nTest))),NSIM))
write.table(tb.df, file="d:/r/2008/ptest/simg/tbn75Numf100.dat")
ti<-paste("Comparison for various snr's with n =",n)
bwplot(TEST~pvalue | SNR, data=tb.df,main=ti)
bwplot(TEST~pvalue | SNR, data=tb.df, subset=TEST!="F",
           scales=list(x=list(relation="free")),main=ti)
ctb<-matrix(numeric(3*nsnr),ncol=3)
for (i in 1:nsnr) {
    out<-t.test(pvalue~TEST, data=tb.df, subset=TEST!="F"&SNR==fsnr[i],paired=TRUE)
    ctb[i,]<-c(out$conf.int,out$p.value)
    }
dimnames(ctb)<-list(snrS, c("lower","upper","pvalue"))
attach(tb.df)
pow<-tapply(pvalue, INDEX=list(SNR, TEST), FUN=function(x) sum(x<0.05)/NSIM)
detach("tb.df")
savews()
```

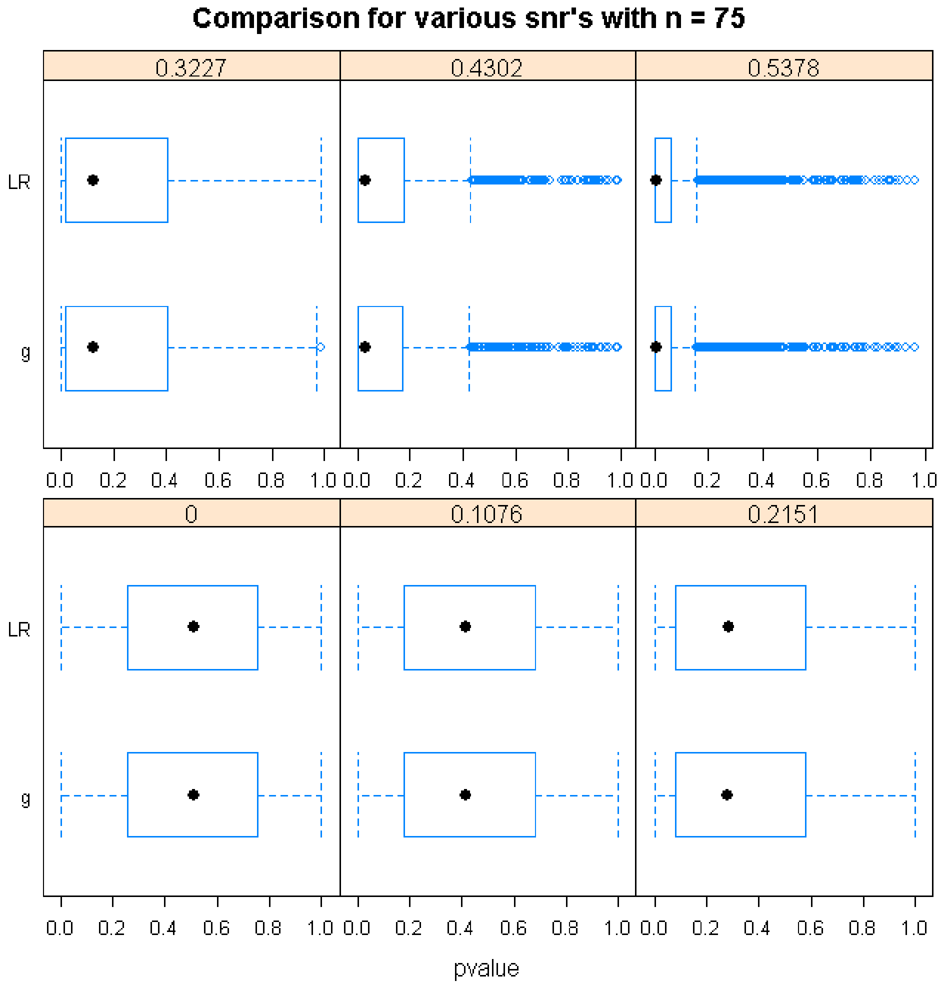
- **Paired t-test for difference in power**

```
> ctb
               lower        upper     pvalue
0      -6.111201e-04 9.577163e-05 0.15285339
0.1076 -6.206880e-04 9.314074e-05 0.14730359
0.2151 -2.040754e-04 4.182540e-04 0.49960838
0.3227  1.005759e-05 5.602854e-04 0.04220725
0.4302 -1.842777e-04 2.940667e-04 0.65252391
0.5378 -1.980256e-04 1.810273e-04 0.92989470
```

- **Power for 5% level test with various snr**

```
           F     g    LR
0      0.049 0.050 0.049
0.1076 0.054 0.091 0.093
0.2151 0.097 0.190 0.195
0.3227 0.173 0.378 0.383
0.4302 0.294 0.583 0.583
0.5378 0.397 0.725 0.725
```


Comparison for various snr's with n = 75
0.3227
0.4302
0.5378
0
0.1076
0.2151
LR
g
F
0.0
0.2
0.4
0.6
0.8
1.0
pvalue

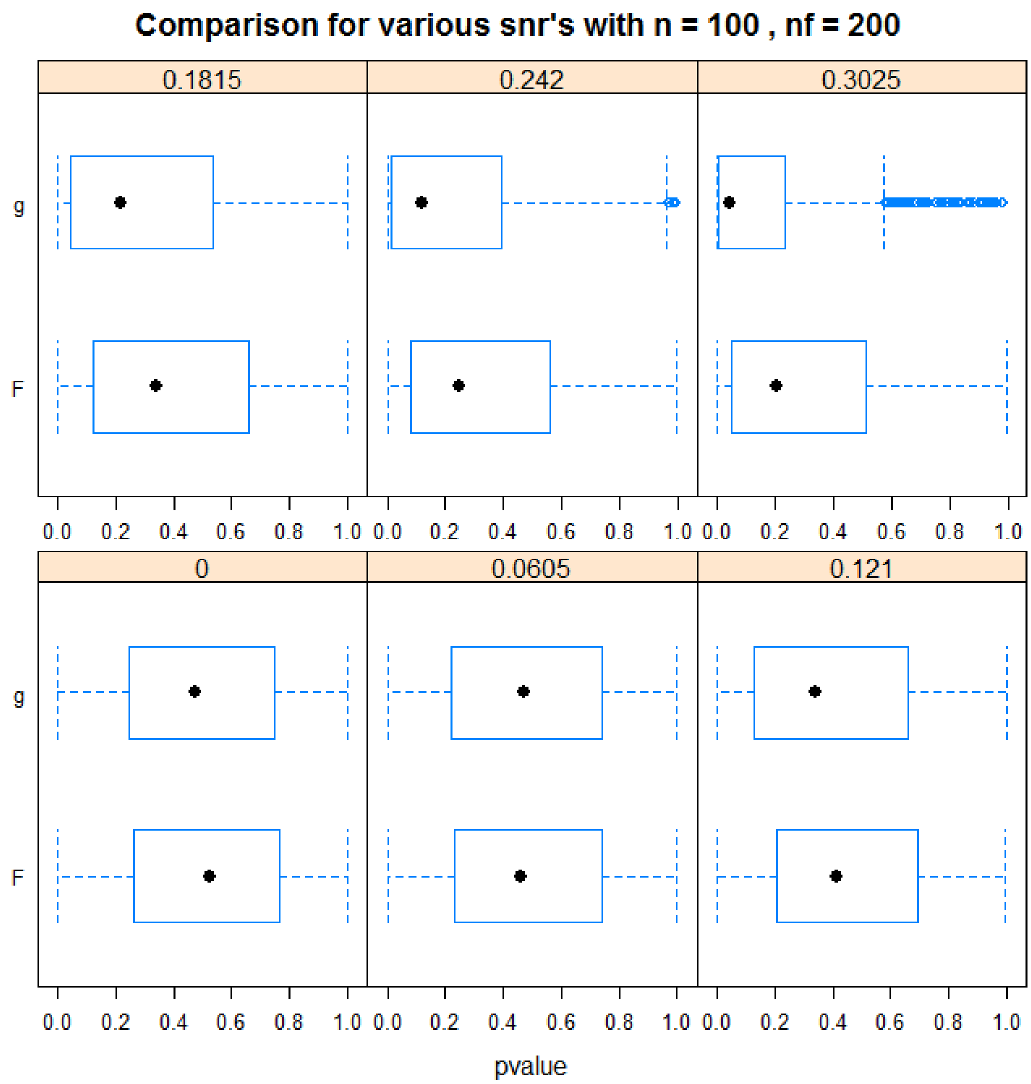


■ Experiment 4. *n* = 100, nf = 200

```
#Comparing Fisher's g statistic and exact Fisher test
start.time<-proc.time()[1]
library(lattice)
n<-100
f<-1.5/n
Numf<-200
NSIM<-1000
snrS<-round(c(0,5,10,15,20,25)*(11/n)^2, 4)
tb<-array(numeric(length(snrS)*NSIM*2),dim=c(length(snrS),2,NSIM))
dimnames(tb)<-list(snrS, c("g", "F"), 1:NSIM)
for (isnr in 1:length(snrS)) {
    snr<-snrS[isnr]
    A<-sqrt(snr)/3
    B<-sqrt(snr-A^2)
    pF<-pg<-numeric(NSIM)
    for (isim in 1:NSIM){
        z<-SimulateHReg(n=n,f=f,A=A,B=B,phi=0, simPlot=FALSE)
        pg[isim]<-PeriodicityTest(z, Numf=Numf, which.test="g", nullDis="b")
        pF[isim]<-FisherGTest(z,UseC=TRUE)$p.value
```

```
    }
    tb[isnr,1,]<-pg
    tb[isnr,2,]<-pF
}
end.time<-proc.time()[1]
total.time<-end.time-start.time
#
#plot results and tabulate t-tests
fsnr<-ordered(snrS)
fTest<-ordered(c("g","F"))
nsnr<-length(snrS)
nTest<-length(fTest)
tb.df<-data.frame(pvalue=c(tb),SNR=rep(fsnr, NSIM*nTest),
           TEST=rep(rep(fTest, rep(rep(nsnr,nTest))),NSIM))
write.table(tb.df, file="d:/r/2008/ptest/simg/tbn100Numf200.dat")
ti<-paste("Comparison for various snr's with n =",n, ", nf =",Numf)
bwplot(TEST~pvalue | SNR, data=tb.df,main=ti, scales=list(x=list(relation="free")))

attach(tb.df)
pow<-tapply(pvalue, INDEX=list(SNR, TEST), FUN=function(x) sum(x<0.05)/NSIM)
detach("tb.df")
savews()
```



- **Power for 5% level test with various snr**

```
> pow
            F     g
0      0.041 0.042
0.0605 0.046 0.072
0.121  0.076 0.139
0.1815 0.149 0.262
0.242  0.196 0.387
0.3025 0.246 0.524
```