

One of the differences between the 7th and 8th editions of Devore's book is the inclusion of the fascinating coverage interval plot in §7.2? It is part of the discussion of why the Agrest-Coull confidence interval for the population proportion is better than the traditional interval.

Both are derived under the null hypothesis that the population proportion of successes p , is some fixed number. A random sample of size n is chosen. Let X be the number of successes in the sample, which has the binomial distribution $X \sim N(n, p)$. The estimator for proportion, which is just the sample mean of n Bernoulli variables, and its standard error are

$$\hat{p} = \frac{X}{n}; \quad \sigma_{\hat{p}} = \sqrt{\frac{p(1-p)}{n}}.$$

By the central limit theorem, and n gets large then the average has approximately the normal distribution $\hat{p} \sim N(p, \sigma_{\hat{p}})$. Suppose we wished to find a level α two-sided confidence interval. Then assuming it is normal, with probability $1 - \alpha$, the standardized variable falls in the range

$$-z_{\alpha/2} < \frac{\hat{p} - p}{\sqrt{\frac{p(1-p)}{n}}} < z_{\alpha/2} \tag{1}$$

The standard interval is found from this by making a further approximation, which lessens its accuracy. It is to approximate that the standard error $\sigma_{\hat{p}}$ by its estimator

$$s_{\hat{p}} = \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

and then solve the inequality (1) with $s_{\hat{p}}$ in place of $\sigma_{\hat{p}}$ yielding the old interval

$$\hat{p} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}. \tag{2}$$

The rule of thumb is that the old interval is a valid interval if $n\hat{p} \geq 10$ and $n(1-\hat{p}) \geq 10$. In that case, the probability that p was captured in this experiment is $1 - \alpha$.

To find their new confidence interval, Argesti and Coull solved (1) for p directly without approximating $\sigma_{\hat{p}}$ to get their interval for p

$$\frac{\hat{p} + \frac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + \frac{z_{\alpha/2}^2}{n}}. \tag{3}$$

Note that under the assumption that the the population proportion is p , we can compute whether or not inequalities (2) or (3) hold. After all, \hat{p} is a binomial random variable, and both intervals boil down to the random variable being between two constants. This is called the coverage probability. I propose to compute it as a function of n and p .

Let us consider the old interval first. The bounds (2) are equivalent to

$$|\hat{p} - p| \leq z_{\alpha/2} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$$

or after squaring

$$(\hat{p} - p)^2 \leq z_{\alpha/2}^2 \frac{\hat{p}(1-\hat{p})}{n}.$$

Combining

$$\left(1 + \frac{z_{\alpha/2}^2}{n}\right) \hat{p}^2 - \left(2p + \frac{z_{\alpha/2}^2}{n}\right) \hat{p} + p^2 < 0$$

This is satisfied if the random variable \hat{p} falls between the roots

$$\frac{p + \frac{z_{\alpha/2}^2}{2n} \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n} + \frac{z_{\alpha/2}^2}{4n^2}}}{1 + \frac{z_{\alpha/2}^2}{n}}. \quad (4)$$

On the other hand, the random variable lies in the Agresti-Coull interval (3) if

$$|\hat{p} - p| \leq z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}$$

or after squaring

$$(\hat{p} - p)^2 \leq z_{\alpha/2}^2 \frac{p(1-p)}{n}.$$

Combining

$$\hat{p}^2 - 2p\hat{p} + p^2 - \frac{z_{\alpha/2}^2}{n} p(1-p) < 0$$

This is satisfied if the random variable \hat{p} falls between the roots

$$p \pm z_{\alpha/2} \sqrt{\frac{p(1-p)}{n}}. \quad (5)$$

The complete symmetry, (2) looks like (5), and (3) like (4) is no surprise, since in deriving the inequalities from (1) we have just reversed the roles of p and \hat{p} .

One more comment is in order. If $X \sim B(n, p)$ is a binomial variable which is discrete, then $B(x, n, p) = P(X \leq x)$, the cumulative distribution function, corresponds to equality. In the case that x is an integer then $P(X < x)$ may be smaller than $P(X \leq x)$. **R** has two functions that give the CDF, namely

$$\begin{aligned} B(x, n, p) &= P(X \leq x) = \text{pbinom}(x, n, p) = 1 - \text{pbinom}(x, n, p, \text{lower.tail} = \text{FALSE}) \\ P(X < x) &= \text{pbinom}(n - x, n, 1 - p, \text{lower.tail} = \text{FALSE}) \end{aligned}$$

Hence

$$\begin{aligned} P(a < X < b) &= P(X < b) - P(X \leq a) \\ &= \text{pbinom}(n - b, n, 1 - p, \text{lower.tail} = \text{FALSE}) - \text{pbinom}(a, n, p). \end{aligned}$$

Finally, if $x(p)$ depends continuously on p , *e.g.*, say if it were one of the interval limits (4) or (5), then $B(x(p), n, p)$ may be a discontinuous function of p since $x(p)$ may be crossing integer values as p varies.

R Session:

R version 2.10.1 (2009-12-14)
Copyright (C) 2009 The R Foundation for Statistical Computing
ISBN 3-900051-07-0

R is free software and comes with ABSOLUTELY NO WARRANTY.
You are welcome to redistribute it under certain conditions.
Type 'license()' or 'licence()' for distribution details.

Natural language support but running in an English locale

R is a collaborative project with many contributors.
Type 'contributors()' for more information and
'citation()' on how to cite R or R packages in publications.

Type 'demo()' for some demos, 'help()' for on-line help, or
'help.start()' for an HTML browser interface to help.
Type 'q()' to quit R.

[R.app GUI 1.31 (5538) powerpc-apple-darwin8.11.1]

[Workspace restored from /Users/andrejstreibergs/.RData]

```
> ##### P(X <= x) VS P(X < x) #####
>
> # Compare P(X <= x) = pbinom(x,n,p) and
> #       P(X < x) = pbinom(n-x, n, 1-p, lower.tail=F)
> # for real values of x as it passes through an integer value.
> p<-10/11;n<-17;c<- 7;x<- (c*11-2):(c*11+2)/11
> x
[1] 6.818182 6.909091 7.000000 7.090909 7.181818
> pbinom(x,n,p)
[1] 2.575798e-08 2.575798e-08 4.105263e-07 4.105263e-07 4.105263e-07
> pbinom(n-x,n,1-p,lower.tail=F)
[1] 2.575798e-08 2.575798e-08 2.575798e-08 4.105263e-07 4.105263e-07
> p<-10/19;n<-23;c<- 11;x<- (c*11-2):(c*11+2)/11
> x
[1] 10.81818 10.90909 11.00000 11.09091 11.18182
> pbinom(x,n,p)
[1] 0.2511192 0.2511192 0.3992274 0.3992274 0.3992274
> pbinom(n-x,n,1-p,lower.tail=F)
[1] 0.2511192 0.2511192 0.2511192 0.3992274 0.3992274
>
> ##### COVERAGE PROBABILITIES OF OLD INTERVALS #####
> alpha <- 0.05
> za<- qnorm(alpha,0,1,lower.tail=F);za
[1] 1.644854
> za2<- qnorm(alpha/2,0,1,lower.tail=F);za2
[1] 1.959964
>
> # The upper / lower limits for X = n * p_hat are a(p) and b(p)
```

```

>
> n <- 20
> c1 <- n/(1+za2^2/n)
> c2 <-n*za2^2/(2*(n+za2^2));c2
[1] 1.611252
> c3 <-n^2*za2/(n+za2^2);c3
[1] 32.88329
> c4 <- za2^2/(4*n^2);c4
[1] 0.002400912
> a <- function(p){c1*p + c2 - c3*sqrt(p*(1-p)/n+c4)}
> b <- function(p){c1*p + c2 + c3*sqrt(p*(1-p)/n+c4)}
> f<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(f(x),0,1,ylim=c(.8,1),main="Old CI Coverage Probability for n=20, alpha=.05")
> abline(h=.95,col=2); abline(v=.5,col=3)
> #M3074CoverageEg1.pdf
>
> n<-50
> c1 <- n/(1+za2^2/n)
> c2 <-n*za2^2/(2*(n+za2^2));c2
[1] 1.78369
> c3 <-n^2*za2/(n+za2^2);c3
[1] 91.00626
> c4 <- za2^2/(4*n^2);c4
[1] 0.0003841459
> a <- function(p){c1*p + c2 - c3*sqrt(p*(1-p)/n+c4)}
> b <- function(p){c1*p + c2 + c3*sqrt(p*(1-p)/n+c4)}
> f<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(f(x),0,1,ylim=c(.8,1),main="Old CI Coverage Probability for n=50, alpha=.05")
> abline(h=.95,col=2);abline(v=c(.2, .8),col=3)
> # M3074CoverageEg2.pdf
>
> n<-100
> c1 <- n/(1+za2^2/n)
> c2 <-n*za2^2/(2*(n+za2^2));c2
[1] 1.849675
> c3 <-n^2*za2/(n+za2^2);c3
[1] 188.7458
> c4 <- za2^2/(4*n^2);c4
[1] 9.603647e-05
> a <- function(p){c1*p + c2 - c3*sqrt(p*(1-p)/n+c4)}
> b <- function(p){c1*p + c2 + c3*sqrt(p*(1-p)/n+c4)}
> f<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(f(x),0,1,ylim=c(.8,1),main="Old CI Coverage Probability for n=100, alpha=.05")
> abline(h=.95,col=2);abline(v=c(.1, .9),col=3)
> # M3074CoverageEg3.pdf

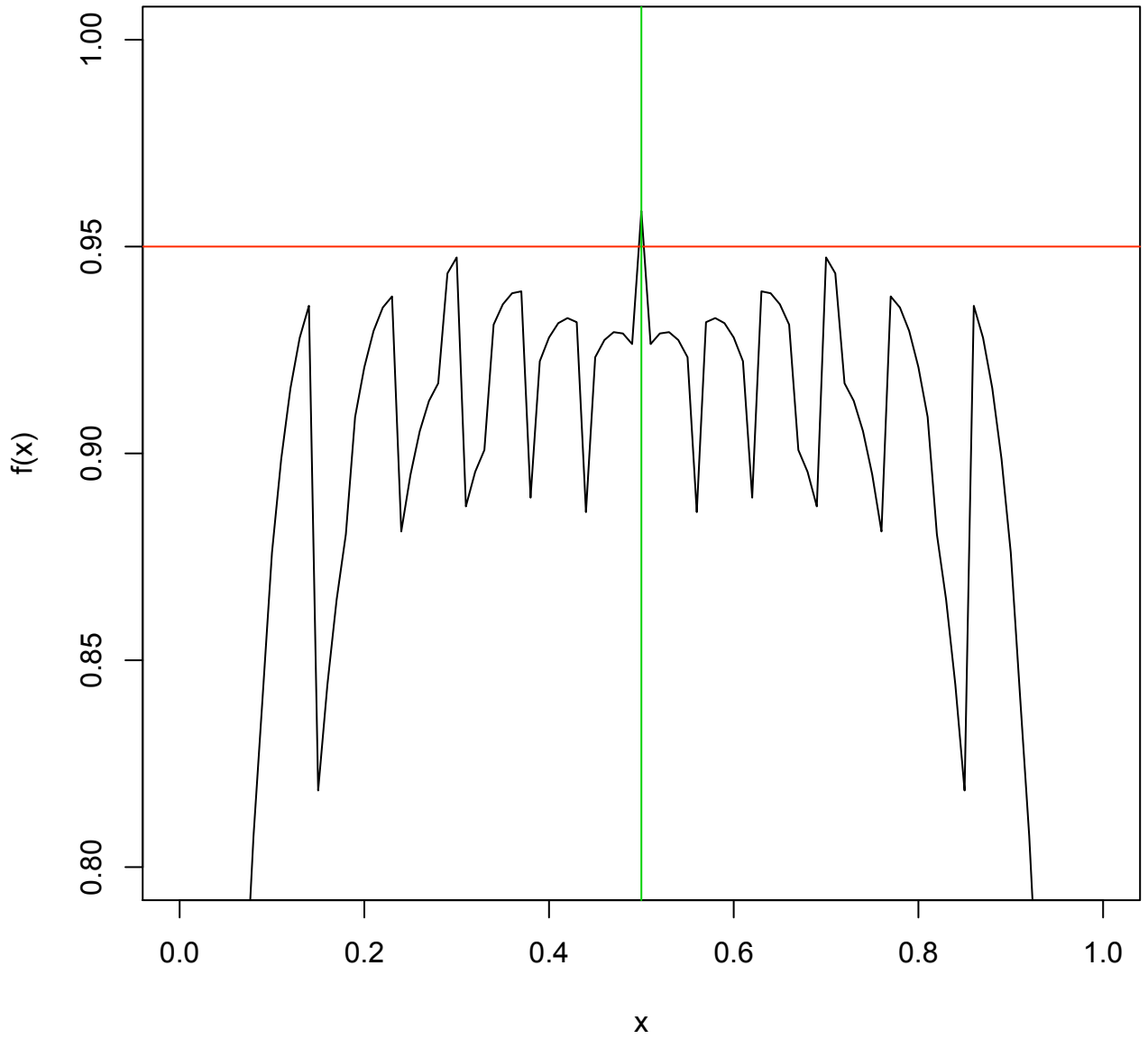
```

```

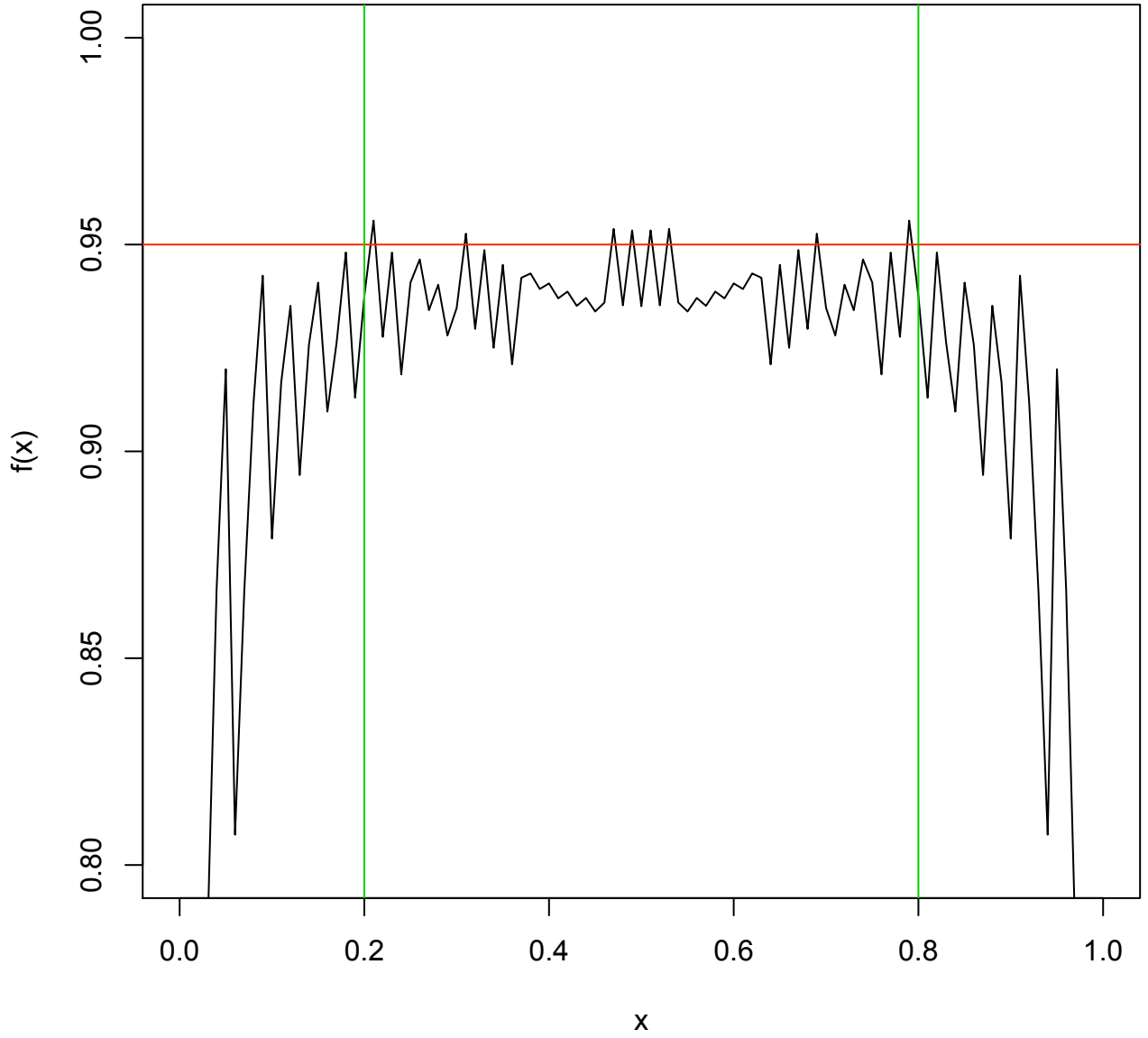
> n<-200
> c1 <- n/(1+za2^2/n)
> c2 <-n*za2^2/(2*(n+za2^2));c2
[1] 1.884533
> c3 <-n^2*za2/(n+za2^2);c3
[1] 384.6056
> c4 <- za2^2/(4*n^2);c4
[1] 2.400912e-05
> a <- function(p){c1*p + c2 - c3*sqrt(p*(1-p)/n+c4)}
> b <- function(p){c1*p + c2 + c3*sqrt(p*(1-p)/n+c4)}
> f<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(f(x),0,1,ylim=c(.8,1),main="Old CI Coverage Probability for n=200, alpha=.05")
> abline(h=.95,col=2);abline(v=c(.05,.95),col=3)
> # M3074CoverageEg3.pdf
>
> n<-400
> c1 <- n/(1+za2^2/n)
> c2 <-n*za2^2/(2*(n+za2^2));c2
[1] 1.902459
> c3 <-n^2*za2/(n+za2^2);c3
[1] 776.5281
> c4 <- za2^2/(4*n^2);c4
[1] 6.00228e-06
> a <- function(p){c1*p + c2 - c3*sqrt(p*(1-p)/n+c4)}
> b <- function(p){c1*p + c2 + c3*sqrt(p*(1-p)/n+c4)}
> f<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(f(x),0,1,ylim=c(.8,1),main="Old CI Coverage Probability for n=400, alpha=.05")
> abline(h=.95,col=2);abline(v=c(1/40,39/40),col=3)
> M3074CoverageEg5.pdf
Error: object 'M3074CoverageEg5.pdf' not found
> # M3074CoverageEg5.pdf
>

```

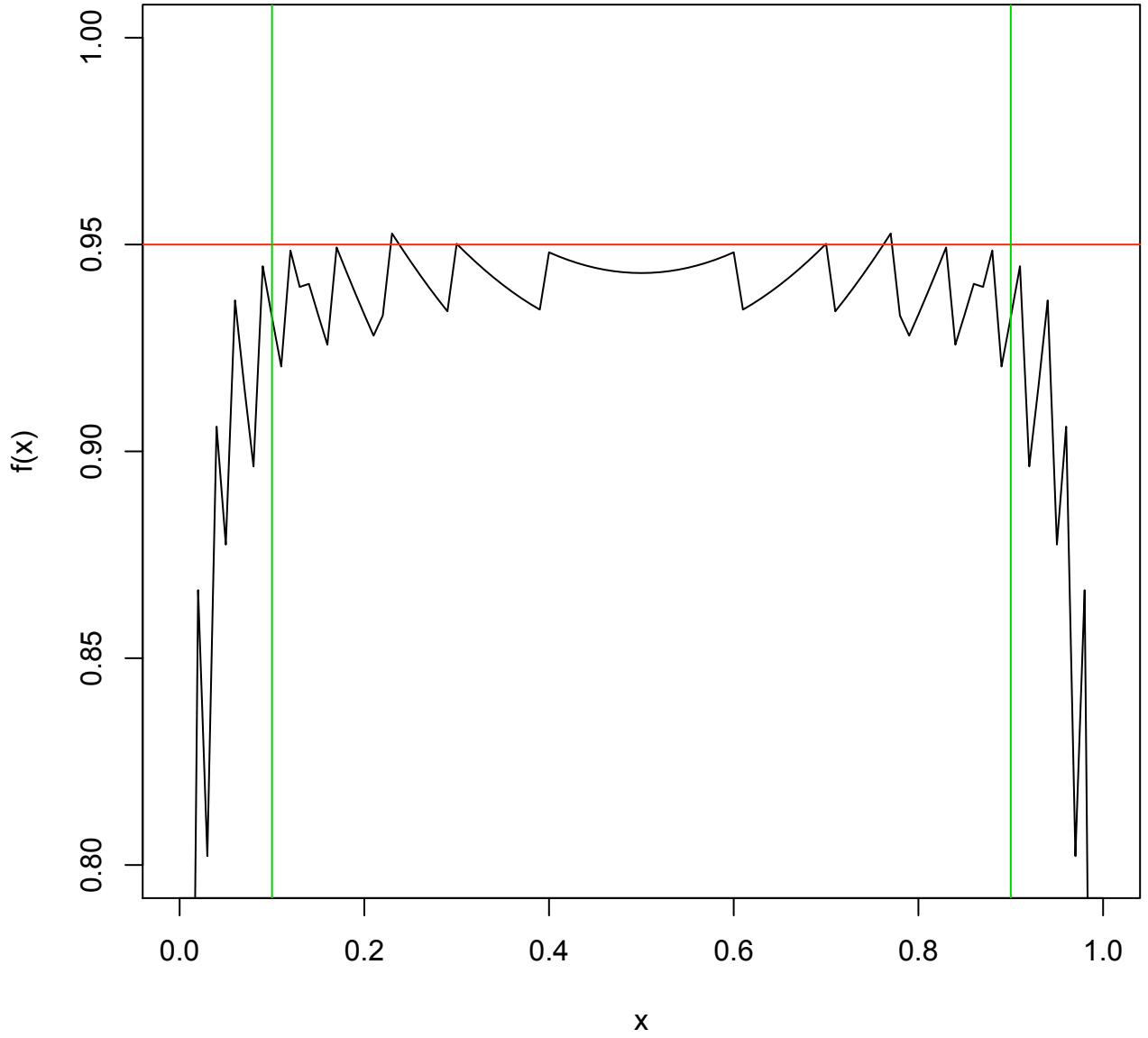
Old CI Coverage Probability for n=20, alpha=.05



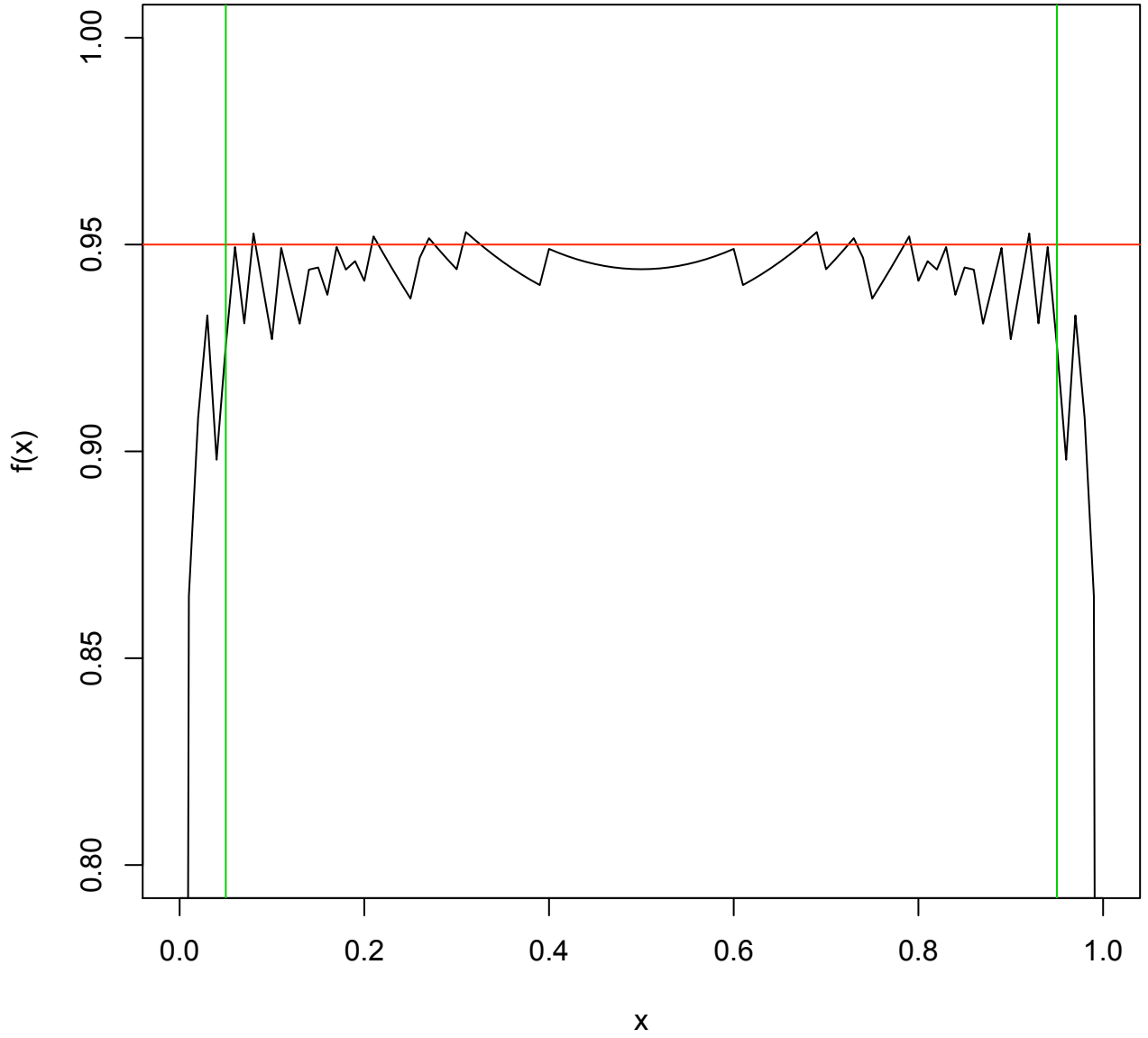
Old CI Coverage Probability for n=50, alpha=.05



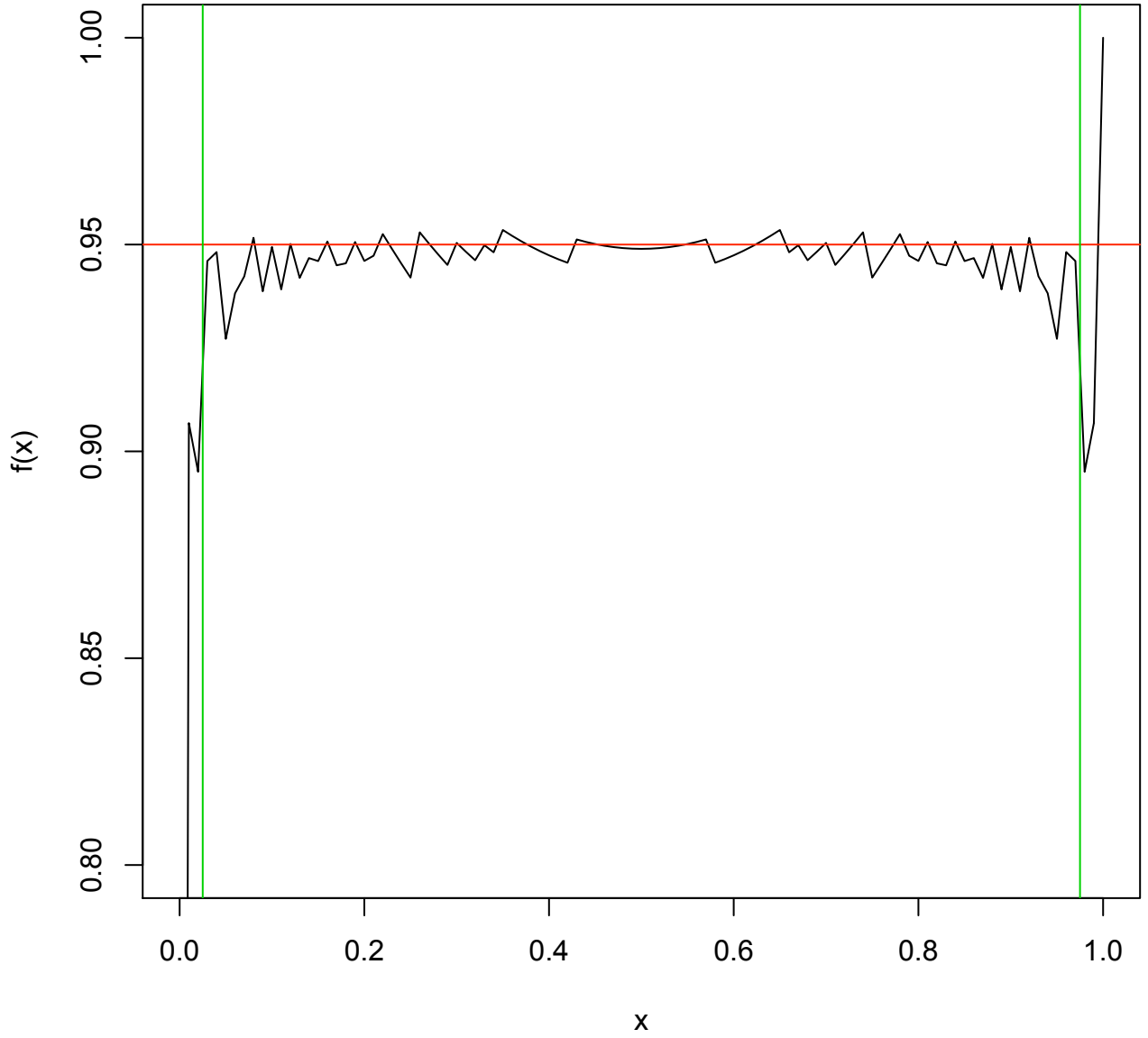
Old CI Coverage Probability for n=100, alpha=.05



Old CI Coverage Probability for n=200, alpha=.05



Old CI Coverage Probability for n=400, alpha=.05



```

> ##### COVERAGE PROBABILITIES OF NEW INTERVALS #####
> n <- 20
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> g<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(g(x),0,1,ylim=c(.8,1),main="New CI Coverage Probability for n=20, alpha=.05")
> abline(h=.95,col=4);abline(v=c(10/n,(n-10)/n),col=5)
> # M3074Coverage8.pdf
>
> n <- 50
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> g<-function(p){pbinom(n-bn(p),n,1-p, lower.tail = FALSE)-pbinom(an(p), n, p)}
> curve(g(x), 0, 1, ylim=c(.8,1),
+ main="New CI Coverage Probability for n=50 alpha=.05",
+ xlab="p", ylab="Probability p-hat Falls Within CI")
> abline(h=.95,col=4);abline(v=c(10/n,(n-10)/n),col=5)
>
> n <- 100
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> g<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(g(x),0,1,ylim=c(.8,1),
+ main="New CI Coverage Probability for n=100 alpha=.05")
> abline(h=.95,col=4);abline(v=c(10/n,(n-10)/n),col=5)
> # M3074Coverage10.pdf
>
>
> n <- 200
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> g<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(g(x),0,1,ylim=c(.8,1),
+ main="New CI Coverage Probability for n=200 alpha=.05")
> abline(h=.95,col=4);abline(v=c(10/n,(n-10)/n),col=5)
> # M3074Coverage11.pdf
>
> n <- 400
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> g<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(g(x), 0, 1, ylim=c(.8,1),
+ main="New CI Coverage Probability for n=200 alpha=05",
+ xlab="p", ylab="Probability p-hat Falls Within CI")
> abline(h=.95,col=4);abline(v=c(10/n,(n-10)/n),col=5)
> # M3074Coverage12.pdf

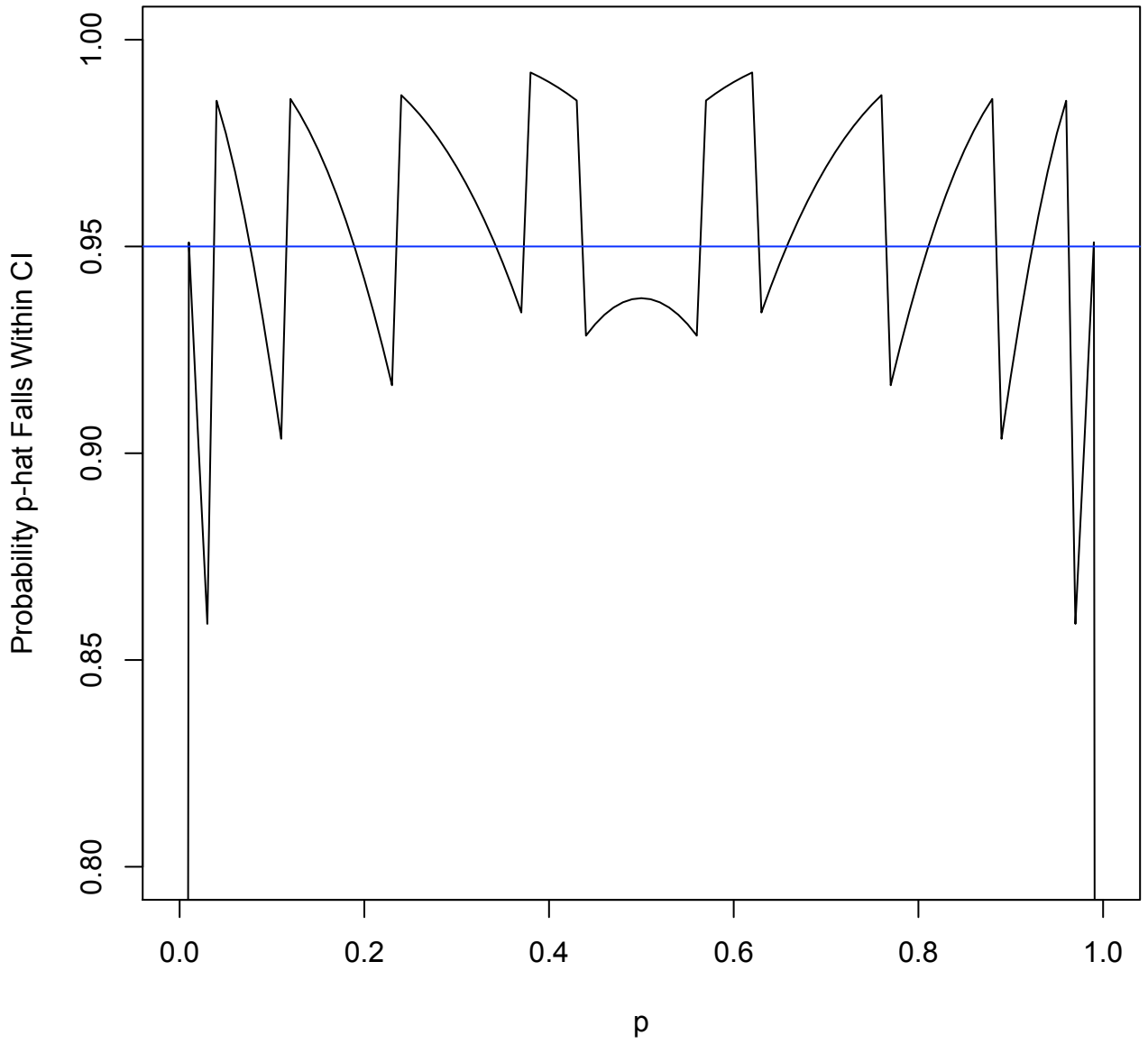
```

```

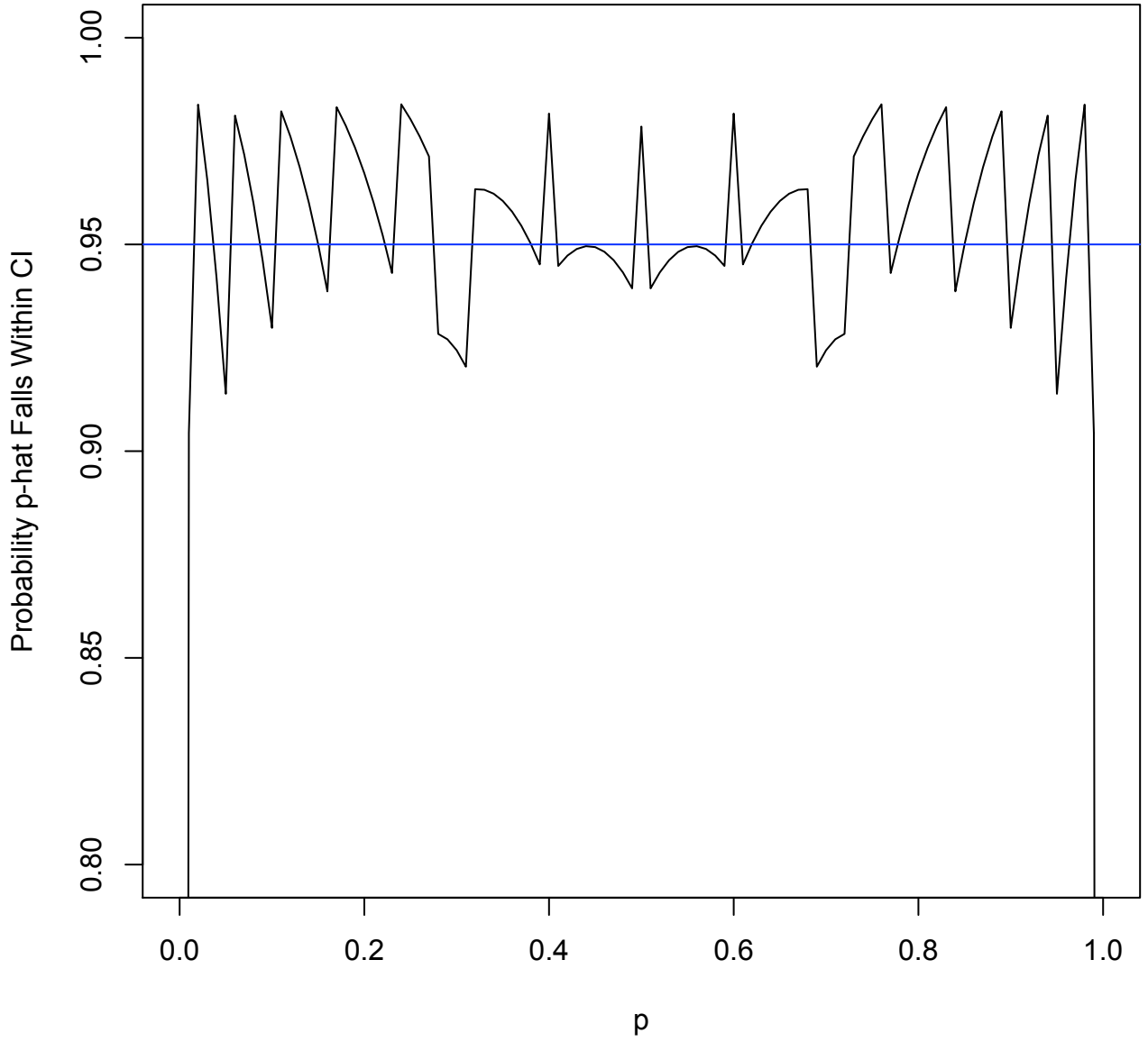
> n <- 10
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> g<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(g(x), 0, 1, ylim=c(.8,1),
+ main="New CI Coverage Probability for n=10 alpha=.05",
+ xlab="p", ylab="Probability p-hat Falls Within CI")
> abline(h=.95,col=4)
> # M3074Coverage7.pdf
>
>
> n <- 5
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> g<-function(p){pbinom(n-b(p),n,1-p, lower.tail = FALSE)-pbinom(a(p), n, p)}
> curve(g(x), 0, 1, ylim=c(.8,1),
+ main="New CI Coverage Probability for n=5 alpha=.05",
+ xlab="p", ylab="Probability p-hat Falls Within CI")
> abline(h=.95,col=4)
> # M3074Coverage6.pdf

```

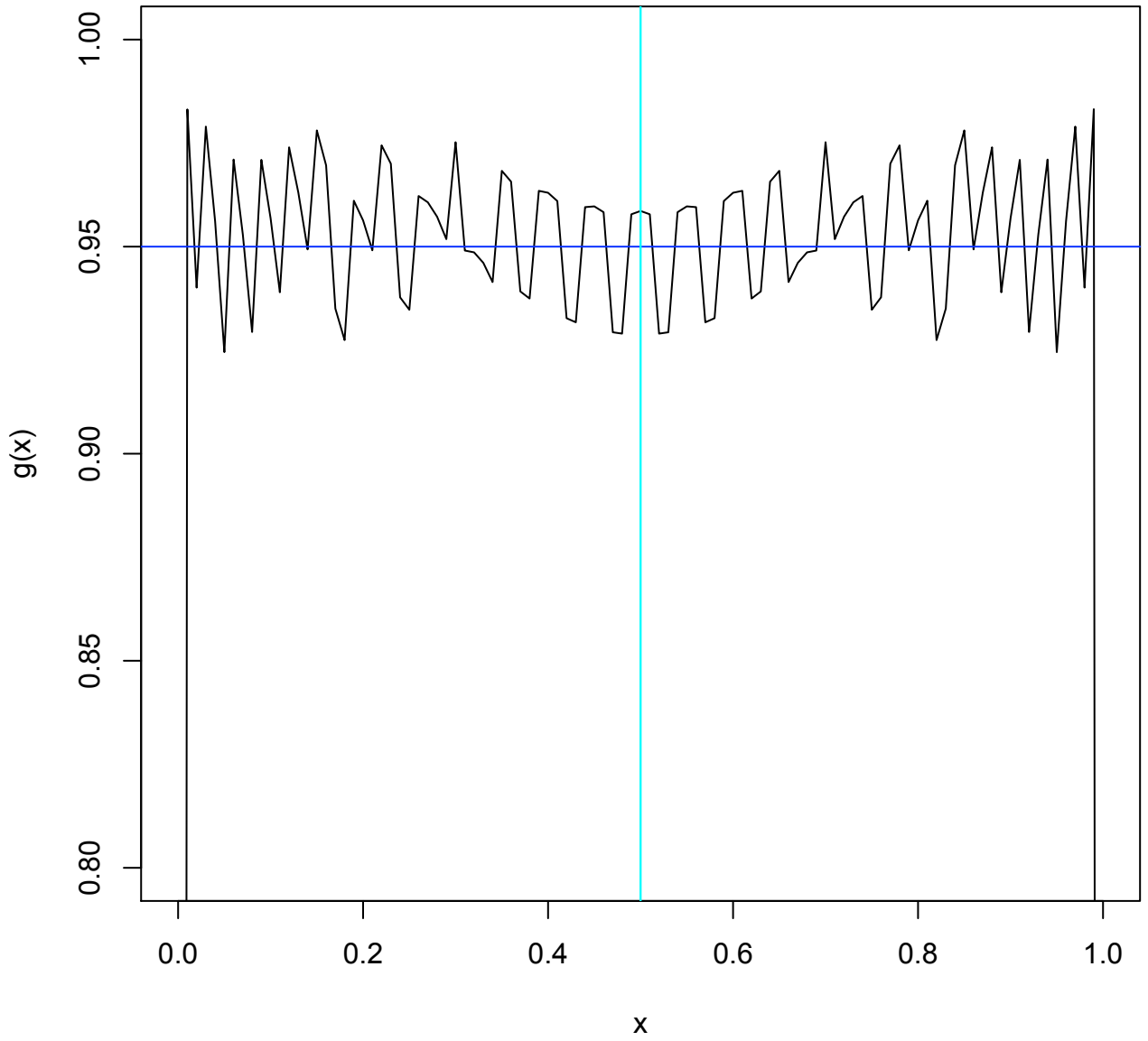
New CI Coverage Probability for n=5 alpha=.05



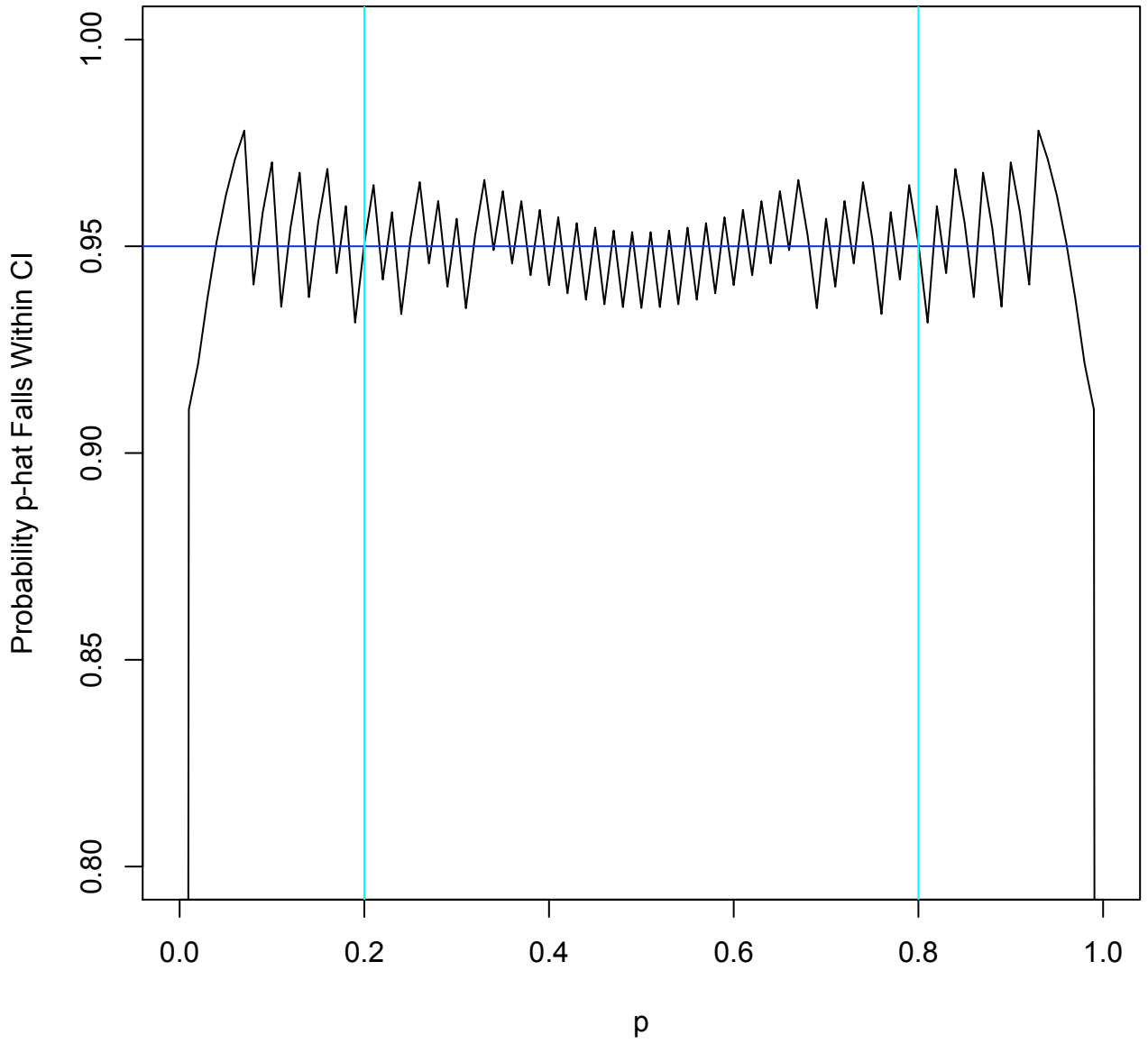
New CI Coverage Probability for n=10 alpha=.05



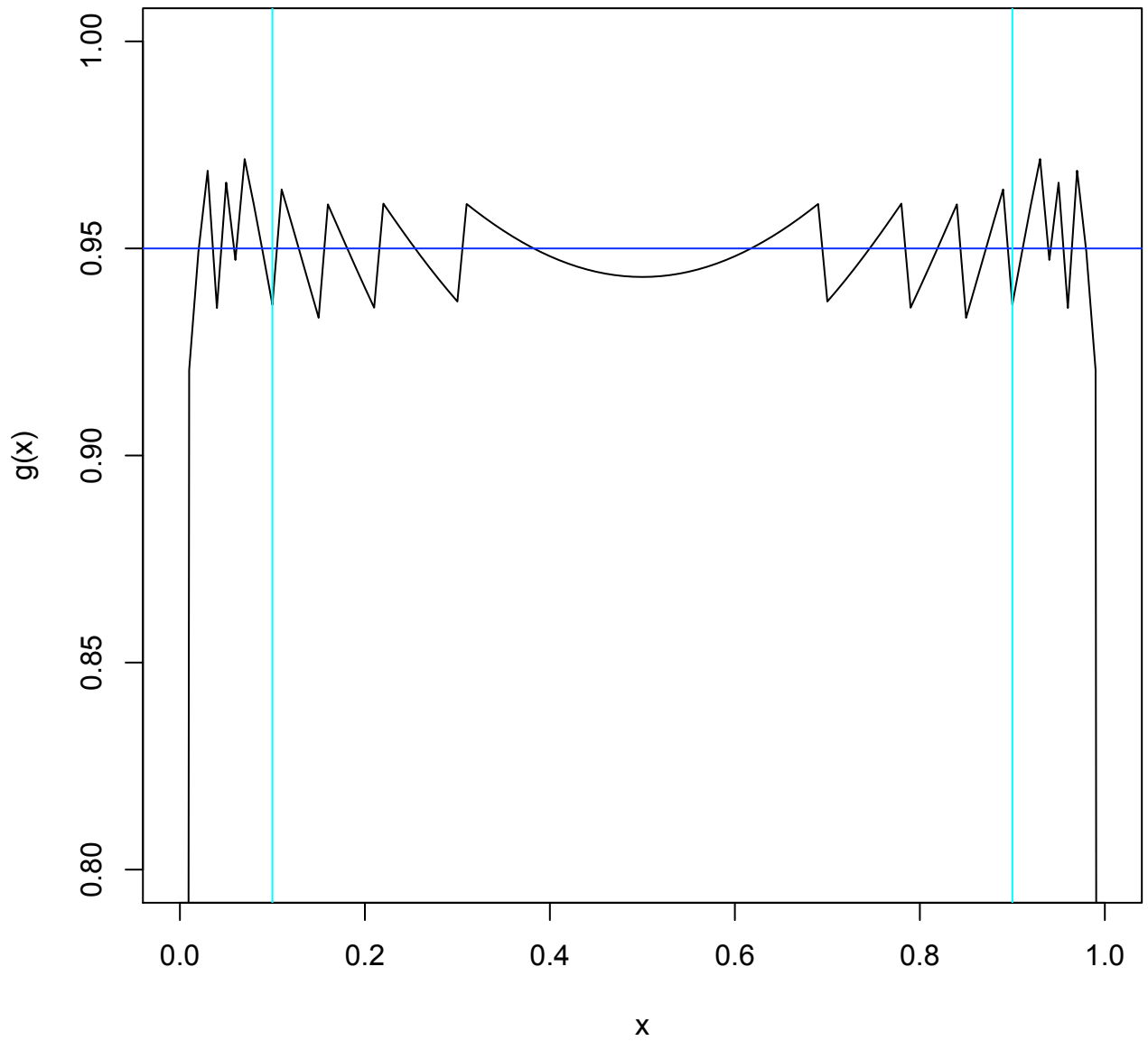
New CI Coverage Probability for n=20, alpha=.05



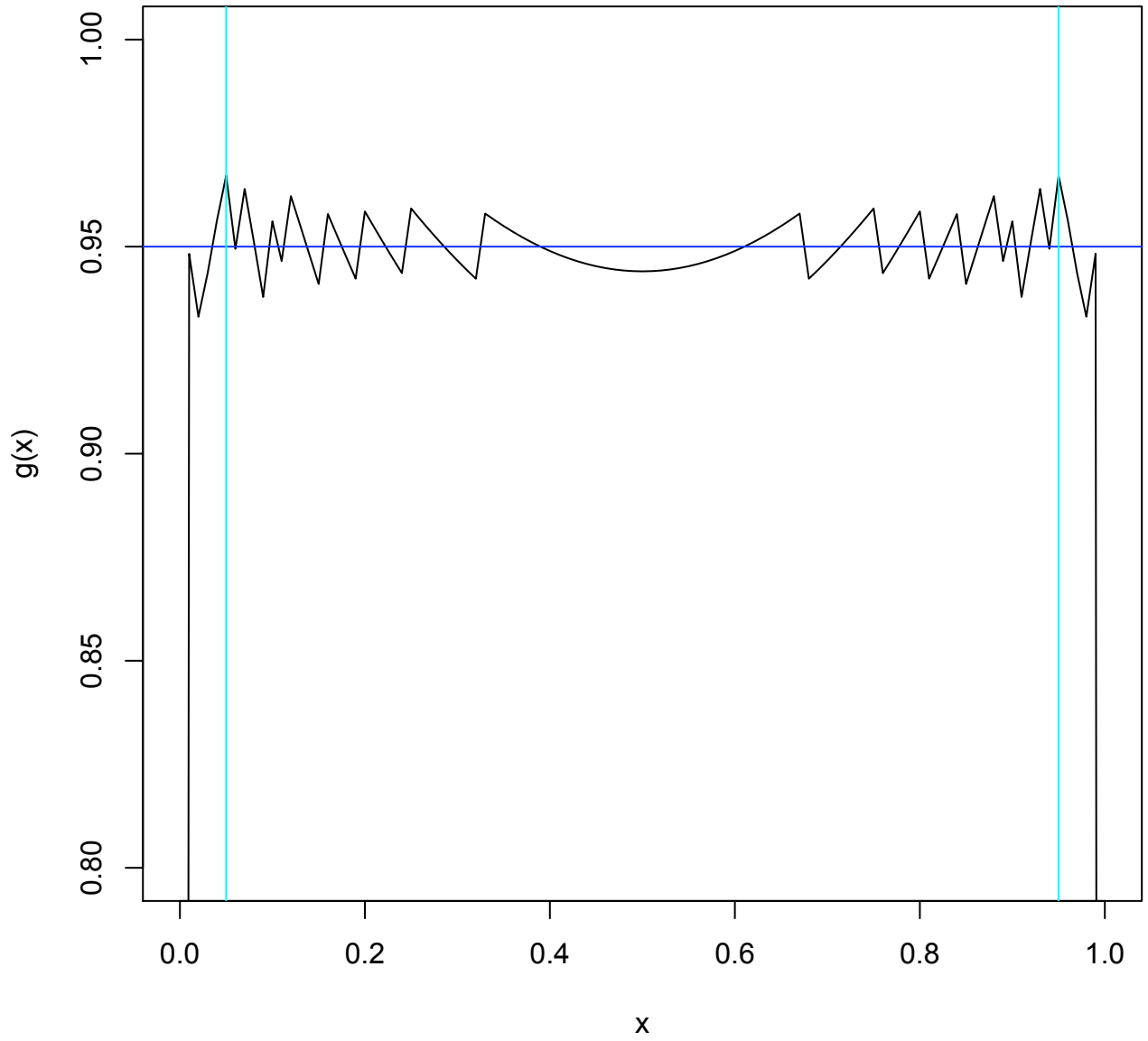
New CI Coverage Probability for n=50 alpha=.05



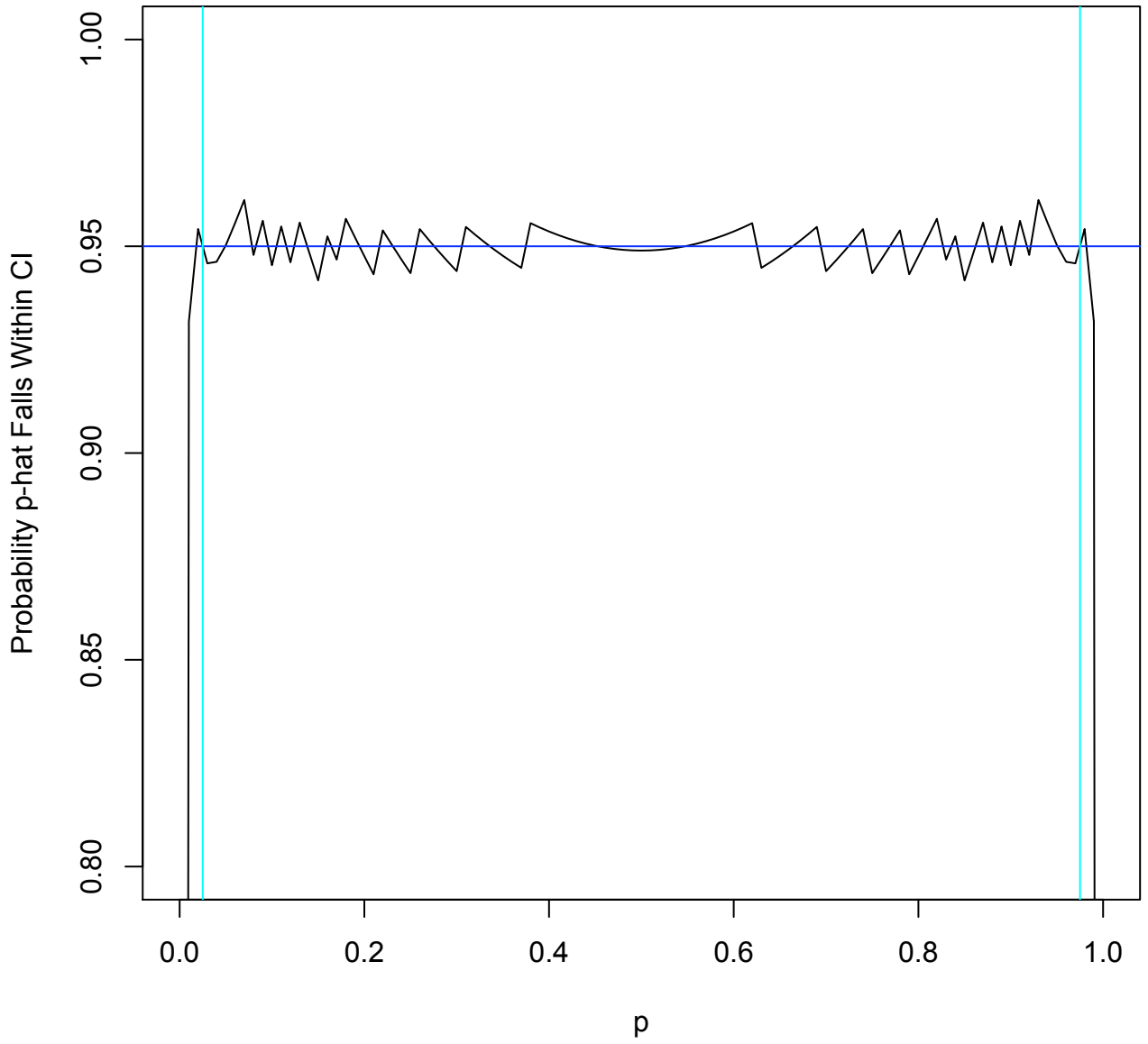
New CI Coverage Probability for n=100 alpha=.05



New CI Coverage Probability for n=200 alpha=.05



New CI Coverage Probability for n=400 alpha=.05



```

> ##### PLOT OF COVERAGE INTERVALS #####
> n <- 20
> an <- function(p){n*p - za2*sqrt(n*p*(1-p))}
> bn <- function(p){n*p + za2*sqrt(n*p*(1-p))}
> c1 <- n/(1+za2^2/n)
> c2 <-n*za2^2/(2*(n+za2^2));c2
[1] 1.611252
> c3 <-n^2*za2/(n+za2^2);c3
[1] 32.88329
> c4 <- za2^2/(4*n^2);c4
[1] 0.002400912
> a <- function(p){c1*p + c2 - c3*sqrt(p*(1-p)/n+c4)}
> b <- function(p){c1*p + c2 + c3*sqrt(p*(1-p)/n+c4)}
> curve(an(x), 0, 1, ylim=c(0,n),
+ main="n=20 Coverage Intervals for Old and New CI's", xlab="p",
+ ylab="Binomial Variable X ~ Bin(n,p)",col=2)
> curve(bn(x),0,1,add=T,col=2)
> curve(a(x),0,1,add=T,col=4)
> curve(b(x),0,1,add=T,col=4)
> abline(h=c(0,20),col="gray")
> abline(v=c(0,1),col="gray")
> legend(0.03,18,legend=c("Old Coverage Interval","New Coverage Interval"),
+ fill=c(4,2))
> # M3074Coverage13.pdf

```

n=20 Coverage Intervals for Old and New CI's

