1

1. Define: Bias of an estimator
   $\mathsf{bias}(\widehat{\theta}) = \mathsf{E}(\widehat{\theta}) - \theta$.

2. Define: Variance of an estimator
   $\mathsf{Var}(\widehat{\theta}) = \mathsf{E}\left(\left(\widehat{\theta} - \mathsf{E}(\widehat{\theta})\right)^2\right)$.

3. Define: Standard error of an estimator
   $\mathrm{SE}(\widehat{\theta}) = \sqrt{\mathsf{Var}(\widehat{\theta})}$.

4. Define: Mean squared error of an estimator
   $\mathrm{MSE}(\widehat{\theta}) = \mathsf{E}\left(\left(\widehat{\theta} - \theta\right)^2\right) = \mathsf{bias}^2(\widehat{\theta}) + \mathsf{Var}(\widehat{\theta})$.

5. Show that the mean squared error can be written as squared bias plus variance.
   See AoS2004, Theorem 6.9. Make sure you define all three quantities. Make sure you understand the proof – especially why the cross term vanishes.

6. Define: Squared loss function
   $L(\widehat{\theta}) = (\widehat{\theta} - \theta)^2$.

7. Define: Risk of an estimator
   $R(\widehat{\theta}) = \mathsf{E}(L(\widehat{\theta}))$. For the squared loss function, risk is same as the MSE of an estimator.

8. Define: Consistency of an estimator
   A point estimator $\widehat{\theta}$ of a parameter θ is said to be consistent if $\widehat{\theta} \xrightarrow{P} \theta$ (AoS2004 Definition 6.7).

9. Define: Asymptotically normal estimator
   AoS2004 Definition 6.12.

10. State and prove: AoS2004 theorem 6.10.
    See AoS2004 Theorem 6.10.

2

1. Let $X_1, \ldots, X_n \sim \text{Uniform}(0, \theta)$ IID. Write any one estimator for θ.

   $\widehat{\theta}_n = 2\overline{X}_n$ (unbiased), or $\widehat{\theta}_n = \max\{X_1, \ldots, X_n\}$ (asymptotically unbiased), or $\widehat{\theta} = (2 + 1/n)\overline{X}_n$ (asymptotically unbiased), $\widehat{\theta}_n = 2 \times \text{median}(X_1, \ldots, X_n)$ (unbiased), or any other reasonable estimator that you can come up with. $\widehat{\theta}_n = \overline{X}_n$ is a badly biased estimator in the present context.

2. Let $X_1, \ldots, X_n \sim \text{Exp}(\lambda)$ IID, and $\widehat{\lambda} = n^{-1} \sum_{i=1}^n X_i$. What is the bias, variance, and MSE of $\widehat{\lambda}$?

   $\mathsf{E}(\widehat{\lambda}) = \mathsf{E}\left(n^{-1} \sum_{i=1}^n X_i\right) = n^{-1}\mathsf{E}\left(\sum_{i=1}^n X_i\right) = n^{-1} \sum_{i=1}^n \mathsf{E}(X_i) = n^{-1} \sum_{i=1}^n \lambda = n^{-1} \times n\lambda = \lambda$. Therefore, $\mathsf{bias}(\widehat{\lambda}) = \mathsf{E}(\widehat{\lambda}) - \lambda = 0$.

   $\mathsf{Var}(\widehat{\lambda}) = \mathsf{Var}\left(n^{-1} \sum_{i=1}^n X_i\right) = n^{-2}\mathsf{Var}\left(\sum_{i=1}^n X_i\right) = n^{-2} \sum_{i=1}^n \mathsf{Var}(X_i) = n^{-2} \sum_{i=1}^n \lambda^2 = n^{-2} \times n\lambda^2 = \lambda^2/n$.

   $\text{MSE}(\widehat{\lambda}) = \mathsf{bias}^2(\widehat{\lambda}) + \mathsf{Var}(\widehat{\lambda}) = \lambda^2/n$.

3. Let $X_1, \ldots, X_n \sim \text{Poisson}(\lambda)$ IID, and $\widehat{\lambda} = n^{-1} \sum_{i=1}^n X_i$. What is the bias, variance, and MSE of $\widehat{\lambda}$? (AoS2004 6.6 exercise 1)

   $\mathsf{E}(\widehat{\lambda}) = \mathsf{E}\left(n^{-1} \sum_{i=1}^n X_i\right) = n^{-1}\mathsf{E}\left(\sum_{i=1}^n X_i\right) = n^{-1} \sum_{i=1}^n \mathsf{E}(X_i) = n^{-1} \sum_{i=1}^n \lambda = n^{-1} \times n\lambda = \lambda$. Therefore, $\mathsf{bias}(\widehat{\lambda}) = \mathsf{E}(\widehat{\lambda}) - \lambda = 0$.

   $\mathsf{Var}(\widehat{\lambda}) = \mathsf{Var}\left(n^{-1} \sum_{i=1}^n X_i\right) = n^{-2}\mathsf{Var}\left(\sum_{i=1}^n X_i\right) = n^{-2} \sum_{i=1}^n \mathsf{Var}(X_i) = n^{-2} \sum_{i=1}^n \lambda = n^{-2} \times n\lambda = \lambda/n$.

   $\text{MSE}(\widehat{\lambda}) = \mathsf{bias}^2(\widehat{\lambda}) + \mathsf{Var}(\widehat{\lambda}) = \lambda/n$.

4. Let $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$ IID, and $\widehat{p} = n^{-1} \sum_{i=1}^n X_i$. What is $\mathsf{Var}(\widehat{p})$?

   $\mathsf{Var}(\widehat{p}) = p(1 - p)/n$. (AoS2004 Example 6.8)

5. Let $X_1, X_2 \sim N(\theta, 1)$ IID. Three estimators for θ have been proposed; viz., (a) $\hat{\theta}_1 = X_1$, (b) $\hat{\theta}_2 = X_2$, and (c) $\hat{\theta}_3 = (X_1 + X_2)/2$. Evaluate the risk functions for each of these three estimators using squared error loss function. Which estimator would you prefer and why?

   (a) $\mathsf{E}(\widehat{\theta}_1) = \mathsf{E}(X_1) = \theta$.
   $\mathsf{bias}(\widehat{\theta}_1) = \mathsf{E}(\widehat{\theta}_1) - \theta = 0$.
   $\mathsf{Var}(\widehat{\theta}_1) = \mathsf{Var}(X_1) = 1$.
   $\text{MSE}(\widehat{\theta}_1) = \mathsf{bias}^2(\widehat{\theta}_1) + \mathsf{Var}(\widehat{\theta}_1) = 1$.

   (b) $\mathsf{E}(\widehat{\theta}_2) = \mathsf{E}(X_2) = \theta$.
   $\mathsf{bias}(\widehat{\theta}_2) = \mathsf{E}(\widehat{\theta}_2) - \theta = 0$.
   $\mathsf{Var}(\widehat{\theta}_2) = \mathsf{Var}(X_2) = 1$.
   $\text{MSE}(\widehat{\theta}_2) = \mathsf{bias}^2(\widehat{\theta}_2) + \mathsf{Var}(\widehat{\theta}_2) = 1$.

   (c) $\mathsf{E}(\widehat{\theta}_3) = \mathsf{E}((X_1 + X_2)/2) = (1/2)\left(\mathsf{E}(X_1) + \mathsf{E}(X_2)\right) = \theta$.
   $\mathsf{bias}(\widehat{\theta}_3) = \mathsf{E}(\widehat{\theta}_3) - \theta = 0$.
   $\mathsf{Var}(\widehat{\theta}_3) = 1/4 \times (\mathsf{Var}(X_1) + \mathsf{Var}(X_2)) = 1/2$.
   $\text{MSE}(\widehat{\theta}_3) = \mathsf{bias}^2(\widehat{\theta}_3) + \mathsf{Var}(\widehat{\theta}_3) = 1/2$.

   All three estimators are unbiased, but $\widehat{\theta}_3$ is preferable because it has the least variance and the least MSE amongst the three.

6. If $X_1, X_2, \ldots, X_n \sim \text{Binomial}(k, p)$ IID, assuming k to be known, design an estimator for p. Compute bias, SE and MSE for this estimator. Construct a 90% confidence interval for p using normal approximation.

   Consider the estimator $\widehat{p} = n^{-1} \sum_{i=1}^n (X_i/k)$:

$\mathsf{E}(\widehat{p}) = \mathsf{E}\left((nk)^{-1}\sum_{i=1}^n X_i\right) = (nk)^{-1}\mathsf{E}\left(\sum_{i=1}^n X_i\right) = (nk)^{-1}\sum_{i=1}^n \mathsf{E}(X_i) = (nk)^{-1}\sum_{i=1}^n(kp) = (nk)^{-1}\times n\times(kp) = p$. Therefore, $\mathsf{bias}(\widehat{p}) = \mathsf{E}(\widehat{p}) - p = 0$.

$\mathsf{Var}(\widehat{p}) = \mathsf{Var}\left((nk)^{-1}\sum_{i=1}^n X_i\right) = (nk)^{-2}\mathsf{Var}\left(\sum_{i=1}^n X_i\right) = (nk)^{-2}\sum_{i=1}^n \mathsf{Var}(X_i) = (nk)^{-2}\sum_{i=1}^n(kp(1-p)) = (nk)^{-2}\times n\times(kp(1-p)) = p(1-p)/nk$.

$\mathrm{SE}(\widehat{p}) = \sqrt{\mathsf{Var}(\widehat{p})} = \sqrt{p(1-p)/nk}$.

$\mathrm{MSE}(\widehat{p}) = \mathsf{bias}^2(\widehat{p}) + \mathsf{Var}(\widehat{p}) = p(1-p)/nk$.

Approximate normal-based confidence interval on p is $\widehat{p}\pm z_{\alpha/2}\widehat{\mathrm{SE}}(\widehat{p}) \equiv \widehat{p}\pm z_{\alpha/2}\sqrt{\widehat{p}(1-\widehat{p})/nk}$. For 90% confidence interval, $\alpha = 0.1$ and $z_{\alpha/2}\approx 1.645$.

7. Suppose $X_1,\dots,X_n$ constitute a random IID sample from a population with mean μ and variance σ^2. Consider estimators of μ of the form

$$\widehat{\mu} = a_1X_1 + a_2X_2 + \ldots + a_nX_n.$$

(a) What condition must be imposed on the coefficients $a_1,\dots,a_n$ so that $\widehat{\mu}$ is unbiased?

Use the definition of bias and the given information about $X_1,\dots,X_n$. Evaluate $\mathsf{bias}(\widehat{\mu})$, and set it to 0. This results in the condition for unbiasedness to be $a_1+\ldots+a_n = 1$.

(b) In addition to the above condition for zero bias, what is the condition for the variance of $\widehat{\mu}$ to be minimum?

$\mathsf{Var}(\widehat{\mu}) = \sigma^2(a_1^2+\ldots+a_n^2)$ because of IID. To find an unbiased, minimum variance estimator $\widehat{\mu}$, we need to minimize $a_1^2+\ldots+a_n^2$ subject to the constraint $a_1+\ldots+a_n = 1$. This can be done using the Lagrange multiplier method. We minimize

$$f(a_1,\dots,a_n;\lambda) = \sum_{i=1}^n a_i^2 + \lambda\left(1-\sum_{i=1}^n a_i\right)$$

with respect to $a_1,\dots,a_n$ and λ, by setting all the partial derivatives to zero:

$$\begin{aligned} 0 &= \frac{\partial f}{\partial a_k} = 2a_k - \lambda \qquad \text{for } k=1,\dots,n \\ 0 &= \frac{\partial f}{\partial \lambda} = 1 - \sum_{i=1}^n a_i. \end{aligned}$$

When we solve these two equations together, we get the solution

$$\begin{aligned} \lambda &= \frac{2}{n} \\ a_k &= \frac{\lambda}{2} = \frac{1}{n} \qquad \text{for } k=1,\dots,n. \end{aligned}$$

So the minimum variance unbiased estimator of the above form turns out to be $\bar{X}_n$.

3

Let $X_1, X_2, \ldots, X_n \sim U(0, \theta)$ IID. Three different estimators for θ, namely,

$$\begin{aligned}
\widehat{\theta}_1 &= \max\{X_1, X_2, \cdots, X_n\} \quad \text{(AoS2004 6.6 Exercise 2)} \\
\widehat{\theta}_2 &= 2\overline{X}_n \quad \text{(AoS2004 6.6 Exercise 3)} \\
\widehat{\theta}_3 &= \tfrac{2+\frac{1}{n}}{n} \textstyle\sum_{i=1}^n X_i
\end{aligned}$$

are proposed.

1. Derive the PDF, mean, standard error, bias, and mean squared error for $\widehat{\theta}_1$.

   Let $Y \equiv \widehat{\theta}_1 = \max\{X_1, X_2, \cdots, X_n\}$. Then

$$\mathsf{P}(Y \le y) = \mathsf{P}(X_1 \le y, \ldots, X_n \le y) = \prod_{i=1}^n \mathsf{P}(X_i \le y) = \prod_{i=1}^n \begin{cases} 0 & y \le 0 \\ y/\theta & 0 \le y \le \theta \\ 1 & y \ge \theta \end{cases},$$

   where we have used the independence of $X_1, X_2, \ldots, X_n$, and that $\mathsf{P}(X_i \le y) = y/\theta$ for $X_i \sim \text{Uniform}(0, \theta)$. Therefore

$$\begin{aligned}
F_Y(y) &= \mathsf{P}(Y \le y) = \begin{cases} 0 & y \le 0 \\ (y/\theta)^n & 0 \le y \le \theta \\ 1 & y \ge \theta \end{cases} \\
f_Y(y) &= \frac{d}{dy} F_Y(y) = \begin{cases} n y^{n-1}/\theta^n & 0 \le y \le \theta \\ 0 & \text{otherwise} \end{cases}.
\end{aligned}$$

   Hence

$$\begin{aligned}
\mathsf{E}(\widehat{\theta}_1) &= \int_0^\theta y f_Y(y) dy = \frac{n}{n+1}\theta \\
\mathsf{Bias}(\widehat{\theta}_1) &= \mathsf{E}(\widehat{\theta}_1) - \theta = -\frac{1}{n+1}\theta \\
\mathsf{Var}(\widehat{\theta}_1) &= \int_0^\theta y^2 f_Y(y) dy - \left(\mathsf{E}(\widehat{\theta}_1)\right)^2 = \theta^2 \left(\frac{n}{n+2} - \left(\frac{n}{n+1} \right)^2 \right) \\
\mathsf{SE}(\widehat{\theta}_1) &= \sqrt{\mathsf{Var}(\widehat{\theta}_1)} = \theta \sqrt{\frac{n}{n+2} - \left(\frac{n}{n+1} \right)^2} \\
\mathsf{MSE}(\widehat{\theta}_1) &= \left(\mathsf{Bias}(\widehat{\theta}_1)\right)^2 + \mathsf{Var}(\widehat{\theta}_1) \quad \text{(Substitute expressions and simplify).}
\end{aligned}$$

   The estimator $\widehat{\theta}_1$ is an example of an *order statistic*: kth order statistic is defined to be kth smallest value amongst $X_1, \ldots, X_n$; $\widehat{\theta}_1$ is, therefore, the nth order statistic. When $X_1, \ldots, X_n \sim \text{Uniform}(0, \theta)$, then the kth order statistics has a Beta distribution; namely, the $B(k, n+1-k)$ distribution.

2. Derive the mean, standard error, bias, and mean squared error for $\widehat{\theta}_2$.

$$\begin{aligned}
\mathsf{E}(\widehat{\theta}_2) &= \mathsf{E}\left(2n^{-1} \sum_{i=1}^n X_i\right) = 2n^{-1} \sum_{i=1}^n \mathsf{E}(X_i) = 2n^{-1} \times n \times (\theta/2) = \theta \\
\mathsf{Bias}(\widehat{\theta}_2) &= \mathsf{E}(\widehat{\theta}_2) - \theta = 0 \\
\mathsf{Var}(\widehat{\theta}_2) &= \mathsf{Var}\left(2n^{-1} \sum_{i=1}^n X_i\right) = 4n^{-2} \sum_{i=1}^n \mathsf{Var}(X_i) = 4n^{-2} \times n \times (\theta^2/12) = \frac{\theta^2}{3n} \\
\mathsf{SE}(\widehat{\theta}_2) &= \sqrt{\mathsf{Var}(\widehat{\theta}_2)} = \frac{\theta}{\sqrt{3n}} \\
\mathsf{MSE}(\widehat{\theta}_2) &= \mathsf{Var}(\widehat{\theta}_2) = \frac{\theta^2}{3n}.
\end{aligned}$$

Incidentally, the distribution of the mean of n Uniform$(0,1)$ RVs is called the Bates Distribution. A related distribution, that of the sum of n Uniform$(0,1)$ RVs, is called Irwin-Hall Distribution.

3. Derive the mean, standard error, bias, and mean squared error for $\widehat{\theta}_3$.

$$
\begin{aligned}
\mathsf{E}(\widehat{\theta}) &= n^{-1}(2+1/n)\sum_{i=1}^{n}\mathsf{E}(X_i) = n^{-1}(2+1/n)n\theta/2 = (2+1/n)\theta/2 \\
\mathsf{bias}(\widehat{\theta}) &= \mathsf{E}(\widehat{\theta}) - \theta = \theta/2n \\
\mathsf{Var}(\widehat{\theta}) &= n^{-2}(2+1/n)^2\sum_{i=1}^{n}\mathsf{Var}(X_i) = n^{-2}(2+1/n)^2 n\theta^2/12 = n^{-1}(2+1/n)^2\theta^2/12 \\
\text{MSE} &= (\mathsf{bias}(\widehat{\theta}))^2 + \mathsf{Var}(\widehat{\theta}) = \theta^2/4n^2 + n^{-1}(2+1/n)^2\theta^2/12.
\end{aligned}
$$

4. Computer Experiment. Draw 1000 samples each of size 30, from $U(0,\theta)$ distribution, with $\theta = 0.1 \times$ (your roll number modulo 100). Apply all three estimators to each of the samples.

```
roll.number <- 0305 # use your own
theta       <- 0.1 * ( roll.number %% 100 )

n <- 30   # data size
m <- 1000 # number of samples

# theta.hat below will be a m X 2 matrix;
# two columns for the two estimators):
theta.hat <- t( sapply( 1:m,
                        function( i )
                         {
                          x <- runif( n, 0, theta )
                          c( max( x ), 2 * mean( x ), ( 2 + 1 / n ) * mean( x ) )
                         }
                     )
              )
```

5. From these samples, estimate the mean and mean squared errors for all three estimators. Are these values close to the values computed in (1)–(3) above?

```
# estimated MSE for the three estimators
mse.hat <- apply( theta.hat, 2, function( x ) { mean( x - theta )^2 } )

# true MSE
mse <- theta^2 * c(
                    1 / ( n + 1 )^2 + n / ( n + 2 ) - ( n / ( n + 1 ) )^2,
                    1 / ( 3 * n ),
                    0.25 / n^2 + ( 2 + 1 / n )^2 / ( 12 * n )
                  )

# difference between true MSE and estimated MSE
mse.table <- cbind( mse, mse.hat, abs( mse - mse.hat ), abs( ( mse - mse.hat ) / mse ) )
colnames( mse.table ) <- c( 'mse', 'mse.hat', 'absdiff', 'reldiff' )
rownames( mse.table ) <- paste( 'theta', 1:ncol( theta.hat ), sep = '' )
print( mse.table )
```

6. Plot histograms of $\frac{\theta-\widehat{\theta}}{\widehat{SE}(\widehat{\theta})}$ for all three estimators.

```
op <- par( mfrow = c( 1, ncol( theta.hat ) ) )
 for ( i in 1:ncol( theta.hat ) )
  hist( ( theta.hat[,i] - theta ) / sd( theta.hat[,i] ),
        breaks = 'FD', col = 'gray', border = 'white', main = paste( 'Estimator', i ),
        xlab = expression( ( hat( theta ) - theta ) / hat( se )( hat( theta ) ) ) )
par( op )
```

Comment on your findings.

4

1. Define: $(1-\alpha)$ confidence interval

   AoS2004 6.3.2

2. What is random? – A confidence interval, or the parameter it traps?

   The confidence interval, because it is a function of the data variables, which are RVs (AoS2004 6.3.2).

3. Define: Coverage of a $(1-\alpha)$ confidence interval

   Coverage of a $(1-\alpha)$ confidence interval on a parameter θ is the least probability with which the confidence interval traps the unknown value of θ. This least probability is $(1-\alpha)$, by definition. (AoS2004 6.3.2).

4. How is a $(1-\alpha)$ confidence interval interpreted?

   AoS2004 6.3.2

5. Derive a confidence interval on the Bernoulli p parameter using Hoeffding's inequality.

   AoS2004 Example 6.15

6. For an asymptotically normal estimator $\widehat{\theta}$ of a parameter θ, what is an approximate/asymptotic $(1-\alpha)$ confidence interval?

   AoS2004 Theorem 6.16. Make sure you understand this theorem and its proof.

7. An estimate $\widehat{\theta}$ of a quantity θ turns out to be π. The estimator $\widehat{\theta}$ is known to be approximately normal with variance π^2. Construct an approximate 95% confidence interval for θ.

   The confidence interval prescription which is based on approximate/asymptotic normality is $\widehat{\theta} \pm z_{\alpha/2}\widehat{\mathsf{se}}(\widehat{\theta})$. Here, $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2)$, and Φ and Φ^{-1} are, respectively, the CDF and the QF of the standard normal distribution. For 95% CI, $\alpha = 0.05$ and $z_{\alpha/2} = \Phi^{-1}(1-\alpha/2) \approx 1.96$. The required CI is, therefore, $\pi \pm 1.96\pi \equiv (-0.96\pi, 2.96\pi)$.

8. Let $Z \sim N(0,1)$, and $z_{\alpha/2} = \Phi^{-1}(1-(\alpha/2))$, where Φ is the $N(0,1)$ CDF.
   What is $P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) =$?

   $P(-z_{\alpha/2} \leq Z \leq z_{\alpha/2}) = 1-\alpha$ (AoS2004 Theorem 6.16).

9. A 95% confidence interval on θ happens to be $(0,5)$. What is the probability that θ is outside this interval?

   5%.

10. What is the $(1-\alpha)$ confidence interval for the Bernoulli p parameter based on asymptotic normality?

    AoS2004 Example 6.17

5

Computer Experiment. Consider $X_1, \ldots, X_n \sim \text{Bernoulli}(p)$ IID, and the two confidence intervals on the parameter p; namely, one based on the Hoeffding inequality, and the other based on asymptotic normality of $\widehat{p} = n^{-1} \sum_{i=1}^{n} X_i$. Fix the value of p, say, $p = 0.3$. Through simulation, find out the empirical coverages of the two confidence interval prescriptions. (Use the repeat-the-same-experiment interpretation of a confidence interval for simulation purposes.) Plot the coverages as functions of n.

```
# estimate through simulation the coverage of the
# two confidence interval prescriptions for Bernoulli(p)

p       <- 0.3                      # Bernoulli p parameter
N       <- c( 5, 10, 50, 100, 500 ) # data sizes
alpha   <- 0.05                     # expected coverage of CI
nrepeat <- 200                      # #replications

zab2 <- qnorm( 1 - 0.5 * alpha )     # Normal
ea2  <- sqrt( log( 2 / alpha ) / 2 ) # Hoeffding

op <- par( mfcol = c( 2, length( N ) ) )

for ( n in N ) # loop over data sizes
 {
  # replicate n coin tosses nrepeat times, and get estimates of p:

  # direct simulation of coin tosses using runif( )
  # p.hat <- replicate( nrepeat, sum( runif( n ) <= p ) / n )

  # short-cut: using the fact that #successes ~ Binomial( n, p )
  p.hat <- rbinom( nrepeat, n, p ) / n

  # for each replicate, compute the two CIs on p:
  ci.n <- sapply( p.hat,
                  function( p.hat )
                   {
                    ci <- p.hat + c( -1, +1 ) * zab2 * sqrt( p.hat * ( 1 - p.hat ) / n )
                    return( pmin( 1, pmax( 0, ci ) ) )
                   }
                )

  ci.h <- sapply( p.hat,
                  function( p.hat )
                   {
                    ci <- p.hat + c( -1, +1 ) * ea2 / sqrt( n )
                    return( pmin( 1, pmax( 0, ci ) ) )
                   }
                )

  # check if each CI contains p:
  contains.n <- apply( ci.n, 2, function( ci ) { ( ci[1] <= p ) && ( ci[2] >= p ) } )
  contains.h <- apply( ci.h, 2, function( ci ) { ( ci[1] <= p ) && ( ci[2] >= p ) } )

  # compute proportion of CIs that contain p:
  coverage.n <- mean( contains.n )
  coverage.h <- mean( contains.h )

  # average CI lengths
  mean.length.n <- mean( apply( ci.n, 2, function( ci ) { diff( ci ) } ) )
  mean.length.h <- mean( apply( ci.h, 2, function( ci ) { diff( ci ) } ) )

  # visualize CIs with reference to p:
  plot.new( ); plot.window( xlim = c( 0, 1 ), ylim = c( 1, nrepeat ) )
  title( main = paste( 'Normal | N =', n, '| Repeat =', nrepeat ),
         xlab  = paste( 'Coverage =', format( coverage.n, digits = 2 ),
                        '| Mean Length =', format( mean.length.n, digits = 2 ) ) )
  axis( 1, at = c( 0, p, 1 ), labels = c( '0', 'p', '1' ) )
  t <- sapply( 1:nrepeat,
               function( i )
                {
                 col = if ( contains.n[i] ) 'green' else 'red'
                 points( ci.n[,i], rep( i, 2 ), pch = 19, cex = 0.2, col = col )
                 lines( ci.n[,i], rep( i, 2 ), col = col )
                }
             )
  abline( v = p, col = 'blue' )
```

```
  plot.new( ); plot.window( xlim = c( 0, 1 ), ylim = c( 1, nrepeat ) )
  title( main = paste( 'Hoeffding | N =', n, '| Repeat =', nrepeat ),
         xlab  = paste( 'Coverage =', format( coverage.h, digits = 2 ),
                        '| Mean Length =', format( mean.length.h, digits = 2 ) ) )
  axis( 1, at = c( 0, p, 1 ), labels = c( '0', 'p', '1' ) )
  t <- sapply( 1:nrepeat,
               function( i )
                {
                 col = if ( contains.h[i] ) 'green' else 'red'
                 points( ci.h[,i], rep( i, 2 ), pch = 19, cex = 0.2, col = col )
                 lines( ci.h[,i], rep( i, 2 ), col = col )
                }
             )
  abline( v = p, col = 'blue' )
 }

par( op )
```