This interval isn’t of the form that you are used to: $\hat{\theta} \pm 2\hat{\sigma}$, for $\hat{\sigma}$ with a factor of $1/\sqrt{n}$.

i. $\text{Var} \left[ \text{median} \right] \approx 1/(4f(\theta)^2n)$ (Cramér, H. (1946) *Mathematical Methods of Statistics*, Princeton U. Press p. 368f. Cramér proves asymptotic normality using an argument that I’m not sure is complete.)

ii. We will learn later on how to estimate this, but it’s harder than the earlier CI rule.

3. Inference for other percentiles

a. Suppose $\theta$ is quantile $\tau \in (0, 1)$ of distribution of iid $X_1, \ldots, X_n$

i. That is, $P[X_j \leq \theta] = \tau$.

b. Generalization of Sign Test:

i. $H_0 : \theta = \theta^o$ vs. $H_A : \theta \neq \theta^o$.

ii. Let $T = \text{number of observations smaller than } \theta$.

iii. Under $H_0$, $T \sim \text{Bin}(n, \tau)$.

iv. $a, b$ so that $\sum_{j=a}^{b-1} \tau^j (1 - \tau)^{n-j} \binom{n}{j} \geq 1 - \alpha$

v. often largest $a$ so that $\sum_{j=0}^{a-1} \tau^j (1 - \tau)^{n-j} \binom{n}{j} < \alpha/2$;

- $a$ is $\alpha/2$ quantile of $\text{Bin}(n, \tau)$ distribution
Lecture 2

- \( a \approx n\tau - \sqrt{n\tau(1-\tau)}z_{\alpha/2} \)

vi. often smallest \( b \) so that \( \sum_{j=b}^{n} \tau^{j} (1-\tau)^{n-j} \binom{n}{j} < \alpha/2 \);

- \( n + 1 - b \) is \( \alpha/2 \) quantile of \( \text{Bin}(n, 1-\tau) \) distribution

- \( b \approx n\tau + \sqrt{n\tau(1-\tau)}z_{\alpha/2} \)

vii. Reject \( H_0 \) if \( T < a \) or \( T \geq b \).

c. Confidence Interval for \( \theta \) is \( (X(a), X(b)) \).

d. Note that confidence level is conservative:

   i. \( P[X(a) \leq \theta \leq X(b)] = 1 - P[X(a) \geq \theta] - P[X\theta \geq X(b)] \geq 1 - \alpha \)

   ii. For any given \( \theta \), inequality is generally strict. [Mark A R]

   [Mark A sas]

H: 2.6

E. Comparing Tests

1. For fixed size, alternative, and power, one with smaller sample size is better.

   a. Make numeric comparison by taking ratio of these two sample sizes: Relative efficiency.

b. Suppose that we have

   i. Parameter \( \theta \) with null value \( \theta^\circ \).
Lecture 2

ii. Alternative is $\theta > \theta^\circ$.

iii. Two families (dependent on sample size) of one-sided test statistics $T_1$ and $T_2$.

iv. Test size $\alpha$, power $1 - \beta$.

c. Compare implied sample sizes

i. Pick $n_1$ sample size for test 1.

ii. Pick $\theta$ so that test $T_1$ of size $\alpha$ has power $1 - \beta$ for alternative $\theta$. Note that $\theta$ depends on $n_1$.

iii. Pick $n_2$ so that test $T_2$ of size $\alpha$ has power $1 - \beta$ for alternative $\theta$.


2. Easy case: Asymptotically normal statistics are compared using standard deviations and derivatives of means under alternative.

a. $T_j \sim N(\mu_j(\theta), \sigma_j^2(\theta)/n_j)$, approximately.

b. Critical value: Find $c$ such that $P_0[T_j \geq c] = \alpha$

c. Since $(T_j - \mu_j(0))/\sigma_j(0)$ is standard normal under the null hypothesis, $P_0\left[\sqrt{n_j}(T_j - \mu_j(0))/\sigma_j(0) \geq z_\alpha\right] = \alpha$.

d. $P_0[T_j \geq \mu_j(0) + \sigma_j(0)z_\alpha/\sqrt{n_j}] = \alpha$.

e. $c = \mu_j(0) + \sigma_j(0)z_\alpha/\sqrt{n_j}$. 
f. Power is \( P_\theta [T_j \geq \mu_j(0) + \sigma_j(0) z_\alpha / \sqrt{n_j}] = 1 - \Phi \left( \sqrt{n_j} (\mu_j(0) + \sigma_j(0) z_\alpha / \sqrt{n_j} - \mu_j(\theta)) / \sigma_j(\theta) \right) \).

g. Approximate with \( \mu_j(\theta) \approx \mu_j(0) + \mu'_j(0) \theta, \sigma_j(\theta) \approx \sigma_j(0) \).

h. Power is \( 1 - \Phi \left( (\sigma_j(0) z_\alpha - \sqrt{n_j} \mu'_j(0) \theta) / \sigma_j(0) \right) = 1 - \Phi \left( z_\alpha - \sqrt{n_j} e_j \theta \right) \) for \( e_j = \mu'_j(0) / \sigma(0) \).

i. Setting power to \( 1 - \beta, z_\alpha - \sqrt{n_j} e_j \theta = z_{1-\beta} \).

j. Solving for \( \theta: \theta = (z_\alpha - z_{1-\beta}) / \sqrt{n_j} e_j \), verifying requirement that \( \theta \) get close to zero.

k. For two tests, with same alternative, \( (z_\alpha - z_{1-\beta}) / \sqrt{n_1 e_1} = (z_\alpha - z_{1-\beta}) / \sqrt{n_2 e_2} \), or \( n_2/n_1 = e_1^2/e_2^2 \).

l. Example: \( T_1 \) is \( t \)-test, \( T_2 \) is sign test, \( \theta \) is median.

i. \( T_1 \sim \) noncentral \( T \), symmetric distribution.

- Simplification: Treat \( T_1 \) as \( Z \) test with separate variances \( \rho \) known.

- \( T_1 \sim N(\theta/\rho, 1/n_1), \mu'_1(0) = 1/\rho, \sigma_1(0) = 1, e_1 = 1/\rho \).

ii. \( T_2 \sim N(\mu_2(\theta), \sigma_2(\theta)^2/n_2) \) for \( \mu_2(\theta) = F(\theta), \sigma_2^2(\theta) = F(\theta)(1 - F(\theta)) \).

iii. \( \mu'_2(0) = f(0), \sigma_2(0) = \frac{1}{2} \),
Lecture 2

iv. Data come from $N(\theta, \rho^2)$:

- $\mu_2'(0) = 1/(\sqrt{2}\pi\rho)$, $\sigma_2(0) = \frac{1}{2}$, $e_2 = \sqrt{2}/\pi/\rho$.
- $n_1/n_2 = (2/\sqrt{2}\pi)^2 = 2/\pi$.

v. Data come from Laplace:

- $\rho = 1$, since Laplace distribution has variance 1, $\mu_1'(0) = 1$, $\sigma_1(0) = 1$, $e_1 = 1$.
- $\mu_2'(0) = 1/\sqrt{2}$, $\sigma_2(0) = \frac{1}{2}$, $e_2 = \sqrt{2}$.
- $n_1/n_2 = (\sqrt{2})^2 = 2$.

vi. Data come from Cauchy: Abuse notation:

- $\rho = \infty$, since Laplace distribution has variance $\infty$, $\mu_1'(0) = 1$, $\sigma_1(0) = \infty$, $e_1 = 0$.
- $\mu_2'(0) = \pi^{-1}$, $\sigma_2(0) = \frac{1}{2}$, $e_2 = 2/\pi$.
- $n_1/n_2 = \infty$.

F. Estimating the CDF of $X_1, \ldots, X_n$

1. identically distributed

2. For $x$ in the range of $X_j$, let $\hat{F}(x)$ be the number of data points less than or equal to $x$, divided by $n$.

3. If observations are independent, $\hat{F}(x) \sim n^{-1} \text{Bin}(n, F(x))$.
   - Confidence interval for $F(x)$ is $\hat{F}(x) \pm$
<table>
<thead>
<tr>
<th>Distribution</th>
<th>( \mu'(0) )</th>
<th>( \sigma(0) )</th>
<th>( e )</th>
<th>( T )-test</th>
<th>Sign test</th>
<th>Relative</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>( 1/\rho )</td>
<td>( 1 )</td>
<td>( 1/\rho )</td>
<td>( 1/(\sqrt{2\pi\rho}) )</td>
<td>( 1/(\sqrt{2\pi\rho}) )</td>
<td>( \sqrt{2/\pi}/\rho )</td>
</tr>
<tr>
<td>Laplace</td>
<td>( 1 )</td>
<td>( 1/\sqrt{2} )</td>
<td>( 1 )</td>
<td>( \sqrt{2}/\rho )</td>
<td>( \sqrt{2}/\rho )</td>
<td>( \sqrt{\pi/2} )</td>
</tr>
<tr>
<td>Cauchy</td>
<td>( \pi^{-1} )</td>
<td>( \infty )</td>
<td>( 0 )</td>
<td>( \sqrt{2} )</td>
<td>( \sqrt{2} )</td>
<td>( 1/\sqrt{2} )</td>
</tr>
</tbody>
</table>

\[ z_{\alpha/2} \sqrt{\hat{F}(x)(1 - \hat{F}(x))}/n, \text{ or exact version.} \]

b. CLT intervals will extend outside \([0, 1]\), which is not reasonable.

i. Use exact, Truncate, or take 542 to learn other approaches.

IV. Two-Sample Testing

A. Independent Samples

1. Data \( Y_1, \ldots, Y_n \) from continuous CDF \( G \)
2. Data \( X_1, \ldots, X_m \) from continuous CDF \( F \)
3. Null hypothesis: distributions are identical: $F(z) = G(z) \forall z$.

4. Usual Approach: Two-sample pooled $t$ test
   
   a. Pooled variance estimate
      
      i. $T = (\bar{Y} - \bar{X}) / \sqrt{s_p^2(1/n + 1/m)}$ for $s_p^2 = 
         \left( \sum_{j=1}^{m}(X_j - \bar{X})^2 + \sum_{j=1}^{n}(Y_j - \bar{Y})^2 \right) / (m + n - 2)$.

5. Nonparametric approach analogous to sign test: Mood’s Median Test.
   
   a. Calculate combined sample median
   
   b. Let $A$ be number of observations from $Y$’s above the combined median
      
      i. Null distribution is hypergeometric, if $F(x) = G(x) \forall x$.
      
      ii. Reduces problem to Fisher’s exact test.

\[
\begin{array}{ccc}
Y & X & \text{Total} \\
> \text{Median} & \quad \quad (m + n)/2 \\
< \text{Median} & A & B = (n + m)/2 - A \\
& m & \quad \quad (m + n)/2 \\
\text{Total} & n & m \\
& \quad \quad m + n \\
\end{array}
\]

iii. Equiv. to score test with $a_j = \begin{cases} 
1 & \text{if } j > (m + n + 1)/2 \\
0 & \text{if } j < (m + n + 1)/2 
\end{cases}$.

6. Advantages:
   
   c. Advantage: simple
   
   d. Disadvantage: low power
      
      i. Consider $X_1, X_2, X_3, Y_1, Y_2, Y_3$ iid under $H_0$. 
ii. We can ignore the ordering of $X_1, X_2, X_3$ among themselves, and similarly with $Y_1, Y_2, Y_3$.

iii. Mood’s test treats $X, Y, X, Y, X, Y$ and $X, Y, X, X, Y, Y$ as having equal evidence against $H_0$, but the second should be treated as having more evidence.


g. The following adaptation for the odd total sample size case seems reasonable:

<table>
<thead>
<tr>
<th></th>
<th>Y</th>
<th>X</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; Median</td>
<td>$B$</td>
<td></td>
<td>$(m + n - 1)/2$</td>
</tr>
<tr>
<td>= Median</td>
<td>$C$</td>
<td></td>
<td>1</td>
</tr>
<tr>
<td>&lt; Median</td>
<td>$A$</td>
<td></td>
<td>$(m + n - 1)/2$</td>
</tr>
<tr>
<td>Total</td>
<td>$n$</td>
<td>$m$</td>
<td>$m + n$</td>
</tr>
</tbody>
</table>

Use as test statistic $B - A$. 
Lecture 2

i. \( P[B - A \geq t] = P[B - A \geq t|C = 0] \left(\frac{m}{m+n}\right)P[B - A \geq t|C = 1] \left(\frac{n}{m+n}\right) \)

h. Called Mood’s Median Test. 

6. Nonparametric approach: Rank sum statistic

a. Assume data are continuous.

b. Rank all of the observations

i. Under null hypothesis, all orderings have equal probabilities

c. Compute \( W = \) sum of ranks in one of the groups (M, perhaps)

d. Calculate expectation and variance

i. \( W = \sum_{j=1}^{m+n} I_{j}j \) for \( I_{j} = \begin{cases} 1 & \text{if subject ranked } j \text{ is from } G \\ 0 & \text{otherwise.} \end{cases} \)

ii. \( E[I_{j}] = \frac{n}{m+n} \)

iii. \( E[W] = \frac{n}{m+n} \sum_{j=1}^{m+n} j = \frac{n(m+n)(m+n+1)}{2(m+n)} = \frac{n(m+n+1)}{2} \)

iv. Variance is harder, since the \( I_{j} \) are not independent.

e. If \( V \) is the sum of ranks for the other group, then

\[ V + W = (n + m)(n + m + 1)/2 \]

f. Compare against normal.

i. Proving this is difficult, since the addends are neither identically distributed nor independent.

: 2.6-2.6.2
g. Alternative formulation:  

\[ W = \sum \text{sum of ranks of } Y \text{'s among whole sample} = \sum_y \#(\text{data points less than or equal to } y) = \sum_y \#(X \text{ values less than or equal to } y) + \sum_y \#(Y \text{ values less than or equal to } y) = U + n(n + 1)/2 \text{ for } U = \sum_{i=1}^{m} \sum_{j=1}^{n} I(X_i < Y_j). \]

i. \( U \) is called Mann-Whitney Statistic. [Mark D R] [Mark D sas]