

But... the formula $s_{CFR} = \sqrt{\frac{CFR(1-CFR)}{N}}$ where $N = D + R$ (i.e. Deaths plus Recoveries) breaks down if $CFR=0$ (0%) or 1 (100%).

In these cases, we find $s_{CFR}=0$ regardless of the value of N .

Furthermore, the **normal approximation** doesn't work well if N is relatively small or if CFR equals or is near 0 or 1.

For example:

If $N=20$ and $CFR = 6/20 = 0.30$ then $s_{CFR} = 0.102$ and the 95% CI is 0.099 to 0.501 (based on the normal approximation)

But

If $N=20$ and $CFR = 1/20 = 0.05$ then $s_{CFR} = 0.049$ and the 95% CI is -0.046 to 0.146 (based on the normal approximation)

This clearly doesn't make sense because CFR can't be negative and furthermore if even just 1 person died, then the CFR can't plausibly be 0.

An alternative is a **likelihood-based binomial 95% confidence interval**

If $N=20$ and **$CFR = 0/20 = 0.0$** , what is the upper limit of the 95% CI?

The answer (obtained numerically) is 0.092. **Making the 95% CI 0.000 to 0.092.**

This is because the log likelihood of the data if $CFR=0.092$ is $20 \cdot \ln(1-0.092) = -1.92$

And the log likelihood of the data if $CFR=0$ is $20 \cdot \ln(1-0) = 0$

Making 2 times the difference equal to $2(0 - -1.92) = 3.84$ (which is the 95th centile of the chi-squared distribution with 1 degree of freedom).

How does this work if CFR doesn't equal 0 or 1?

Similarly!

If $N=20$ and **$CFR = 1/20 = 0.05$** , what are the limits of the 95% CI?

The answers (obtained numerically) are 0.0029 and 0.202. **Making the 95% CI 0.0029 to 0.202.**

This is because the log likelihood of the data if $CFR=0.202$ is $1 \cdot \ln(0.202) + 19 \cdot \ln(1-0.202) = -5.89$

And the log likelihood of the data if $CFR=0.0029$ is $1 \cdot \ln(0.0029) + 19 \cdot \ln(1-0.0029) = -5.89$

And the log likelihood of the data if $CFR=0.05$ is $1 \cdot \ln(.05) + 19 \cdot \ln(1-0.5) = -3.97$

Making 2 times the difference equal to $2(-3.97 - -5.89) = 3.84$ (which is the 95th centile of the chi-squared distribution with 1 degree of freedom).

Another alternative is the **exact binomial 95% confidence interval**.

Here the equations for the lower and upper bounds of the 95% confidence interval (p_{LB} and p_{UB} , respectively) are:

$$\sum_{k=0}^r \binom{N}{k} p_{UB}^k (1 - p_{UB})^{(N-k)} = 0.025$$

$$\sum_{k=r}^N \binom{N}{k} p_{LB}^r (1 - p_{LB})^{(N-k)} = 0.025$$

where r is the number of deaths and N is the sample size.

For N=20 and CFR = 0/20 = 0.0, the exact 95% CI is **(0.0, 0.1684)**

For N=20 and CFR = 1/20 = 0.05, the exact 95% CI is **(0.0013, 0.2487)**