Approximate CI's for mean, μ , of a Poisson distribution, based on 5 different approximations to Poisson tail areas

(1) Wilson/Hilferty approxn. to Chi-square quantiles.

[helpful when appropriate Chi-square quantiles not readily available]

This approximation , which has high accuracy for **c** > 10, uses z , the normal standardized variate corresponding to α, e.g.. z = 1.645 for α = 0.05,1.96 for α = 0.025, etc.

$$\mu_{LOWER} = (\mathbf{c}) \{ 1 - (9\mathbf{c})^{-1} - z \ (9\mathbf{c})^{-1/2} \}^3$$

$$\mu_{UPPER} = (\mathbf{c+1}) \{ 1 - (9[\mathbf{c+1}])^{-1} + z \ (9[\mathbf{c+1}])^{-1/2} \}^3$$

Note1: Rothman[2002], page 134, provides an adaptation from "D. Byar, unpublished" in which he makes a further approximation, using the average (c+0.5) for both the lower an upper limits, rather than the more accurate c for the lower and c+1 for the upper limit. This is called method 1' below. JH is surprised at Rothman's eagerness to save a few keystrokes on his calculator, and at his reference to an unpublished source, rather than the 1931 publication of Wilson & Hilferty. Full W-H citation, and evaluation of the above equation, in Liddell's "Simple exact analysis of the standardized mortality ratio" in J Epi and Comm. Health 37 85-88, 1984 available on 626 website.

Note2: Rothman uses the CI for the expected numerator of a Rate. {e.g.s below focus on number in same sized study, not rate *per se*.

(2) Square-root transformation of Poisson variable.

With μ large enough, $\sqrt{\mathbf{c}}$ is approximately Gaussian with mean $\sqrt{\mu}$ and variance 1/4 or SD 1/2 (the variance and SD are thus independent of $\sqrt{\mu}$) .

This leads to (see ref. (3)):

$$\mu_{LOWER,UPPER} = \mathbf{c} \ \mathbf{-/+} \ z\,(\mathbf{c})^{1/2} + {}^{1}/_{4}(z)^2$$

This simpler formula is accurate when c > 100 or so.

(3) 1st Principles CI from c ~ Gaussian(μ , SD = $\sqrt{\mu}$)

Obtained by solving the two equations:

$$c = \mu_{LOWER} + z\sqrt{\mu_{LOWER}} \ ; \ c = \mu_{UPPER} - z\sqrt{\mu_{UPPER}}$$

to give

$$\mu_{LOWER,UPPER} = \left(\sqrt{c + z^2/4} \ -/+ \ z/2\right)^2$$

"First Principles" : it recognizes that Poisson variance is different (smaller) at $\mu = \mu_{LOWER}$ than at $\mu = \mu_{UPPER}$.

(4) (Naive) CI based on c ~ Gaussian(μ , $\hat{SD} = \sqrt{c}$).

If really lazy, or don't care about principles or accuracy, or if c is large (3 digits) might solve

$$c = \mu_{LOWER} + z\sqrt{c} \ ; \ c = \mu_{UPPER} - z\sqrt{c}$$

to give

$$\mu_{LOWER,UPPER} = c \ -/+ \ z\sqrt{\mathbf{c}}$$

Accuracy of 5 approximations (95% CI's) in 5 eg's

Method	c = 3*	c = 6	c = 33**	c=78***	c=100
Exact	(2.48,35.1)	(2.20,13.1)	(22.7,46.3)	(61.7,97.3)	(81,122)
(1)	(2.41,35.1)	(2.19,13.1)	(22.7,46.3)	(61.7,97.3)	(81,121)
(1')	(3.32,32.0)	(2.49,13.4)	(23.1,45.8)	(62.1,96.8)	(82,121)
(2)	(2.26,29.4)	(2.16,11.8)	(22.7,45.2)	(61.7,96.3)	(81,121)
(3)	(4.08,35.3)	(2.75,13.1)	(23.5,46.3)	(62.5,97.3)	(82,122)
(4)	(–1.6,25.6)	(1.20,10.8)	(21.7,44.3)	(60.7,95.3)	(80,120)

* Rothman2002 p134 "3 cases in 2500 PY; pt. est. of Rate:12 per 10 000PY
Focus: No. per 10000PY (Rate) *rather than* on ave. No. in 2500PY
Focus for c=6, 33, 78 & 100: ave. No in same-size study (no den. given)

** No. of cancers among females and ***overall in Alberta SourGas study