1. z-CI for μ (equal probability with replacement sampling):

$\boxed{\bar{x} \pm z \frac{s}{\sqrt{n}}}$ (z = 1, 1.96, 3.09 for 68%, 95%, 99% confidence)

P(μ is covered by $\boxed{\bar{x} \pm z \frac{s}{\sqrt{n}}}$) ~ P($|Z| < z$), n large.

2. Hybrid z-CI for μ (eq-prob with replacement sampling):

Desired hybrid z-CI half width W > 0 is specified in advance.
Preliminary sample size must be suitable for applying z-CI.
Determine final sample size

$n_{final} = \left(z\, s_{prelim} / W\right)^2$ (z = 1, 1.96, for 68%, 95%, etc.)

$\boxed{\bar{x}_{final} \pm W}$ (but for rounding, $W = z\, s_{prelim} / \sqrt{n_{final}}$).

P(μ **in** $\boxed{\bar{x}_{final} \pm W}$) ~ P($|Z| < z$), n_{prelim} large.

If $n_{prelim} \geq n_{final}$ just use z-CI from preliminary sample.

3. z-CI for μ (equal probability withOUT replacement sampling):

$\boxed{\bar{x} \pm z \frac{s}{\sqrt{n}} FPC}$ with $FPC = \sqrt{(N - n)/(N - 1)}$

P(μ is covered by $\boxed{\bar{x} \pm z \frac{s}{\sqrt{n}} FPC}$) ~ P($|Z| < z$), n, N-n large.

4. t-CI for μ (sampling from a NORMAL x distribution):

$\boxed{\bar{x} \pm t_{\alpha, df} \frac{s}{\sqrt{n}}}$ ($t_{.025, \infty} = 1.96$, $t_{.025, 2} = 4.303$, $df = n - 1$)

P(μ is covered by $\boxed{\bar{x} \pm t_{\alpha, df} \frac{s}{\sqrt{n}}}$) = P($|T_{df}| < t_{\alpha, df}$)

ideally (but for approximations in calculations) for n > 1.

5. z-CI for μ_x - μ_y (paired data, utilizing difference scores):

$\boxed{\bar{d} \pm z \frac{s_d}{\sqrt{n}}}$ (z = 1, 1.96, 3.09 for 68%, 95%, 99% confidence)

P($\mu_d = \mu_x$-μ_y is covered by $\boxed{\bar{d} \pm z \frac{s_d}{\sqrt{n}}}$) ~ P($|Z| < z$), n large.

(n is the number of pairs, each of which has scores (x, y)).

6. z-CI for μ_x-μ_y (utilizing UNpaired data):

$$\boxed{(\bar{x} - \bar{y}) \pm z \sqrt{s_x^2 / n_x \oplus s_y^2 / n_y}}$$

P(μ_x-μ_y is covered by $\boxed{(\bar{x} - \bar{y}) \pm z \sqrt{s_x^2 / n_x \oplus s_y^2 / n_y}}$)

~ P($|Z| < z$), n_x, n_y large.

7. z-CI for μ_x (utilizing KNOWN strata population rates):

$$\left(\sum_i \mathbf{W_i}\, \overline{\mathbf{x}_i}\right) \pm z\sqrt{\sum_i W_i^2\, s_i^2 \,/\, n_i}$$

P(μ_x is covered by $\left(\sum_i \mathbf{W_i}\, \overline{\mathbf{x}_i}\right) \pm z\sqrt{\sum_i W_i^2\, s_i^2 \,/\, n_i}$)

~ P($|Z| < z$), n_x large.

Weights W_i must be the KNOWN fractions of the population in each stratum i. For example W_1 could denote the KNOWN fraction of males in the POPULATION, $\overline{x_1}$ denoting the SAMPLE mean age of men.

8. z-CI for μ_x (utilizing KNOWN population mean μ_y):

$$\left(\overline{\mathbf{x}} + \left(\mu_{\mathbf{y}} - \overline{\mathbf{y}}\right)\frac{\overline{\mathbf{xy}} - \overline{\mathbf{x}}\,\overline{\mathbf{y}}}{\overline{\mathbf{y}^2} - \overline{\mathbf{y}}^2}\right) \pm Z\,\frac{s_x}{\sqrt{n}}\sqrt{1 - r^2}$$

P(μ_x is covered by $\left(\overline{\mathbf{x}} + \left(\mu_{\mathbf{y}} - \overline{\mathbf{y}}\right)\frac{\overline{\mathbf{xy}} - \overline{\mathbf{x}}\,\overline{\mathbf{y}}}{\overline{\mathbf{y}^2} - \overline{\mathbf{y}}^2}\right) \pm Z\,\frac{s_x}{\sqrt{n}}\sqrt{1 - r^2}$)

~ P($|Z| < z$), n large.

Variable y must be gathered as data paired with x and the population mean μ_y **must be known. For example, we are sampling business owners to estimate μ_x = population mean loss in business compared with last year.** On the supposition that last year's tax y paid by the business may be correlated with x, and thus offer improved estimation for μ_x, **we decide to ask each owner also for their tax paid last year.** The average tax μ_y for the population of all businesses is a number we can come up with. Our z-CI is narrower by the factor $\sqrt{1 - r^2} < 1$, where r denotes the sample correlation (of x with y) defined by

$$r = \frac{\overline{xy} - \overline{x}\,\overline{y}}{\sqrt{\overline{x^2} - \overline{x}^2}\,\sqrt{\overline{y^2} - \overline{y}^2}}$$

9. Chi-Square statistic, df, P-value.

$$\chi^2 = \sum_{\text{cells}} \frac{(O-E)^2}{E}$$

$$\text{df} = \#\text{cells} - 1 - \#(\text{estimations needed to determine expected counts from data})$$

P-value = P(χ^2 > chi-square statistic as seen from data)

For example, the model

AA	Aa	aa	
p^2	2p(1-p)	$(1-p)^2$	for p = $\frac{2\,\#\,\text{AA} + \#\,\text{Aa in population}}{2\,\#\,\text{population}}$

with data

16	8	6	total of 30 samples

we estimate p by

$$\hat{p} = \frac{2\,\#\,\text{AA} + \#\,\text{Aa in sample}}{2 \times 30} = \frac{2\ 16 + 8}{2 \times 30} = \frac{40}{60} = \frac{2}{3}$$

Pro-rating 30 observations in accordance with this $\hat{p}$ we get

E for AA = 30 $\hat{p}^2$ = 30 $(2/3)^2$ ~ 13.333333

E for Aa = 30 2 $\hat{p}$ (1-$\hat{p}$) = 30 2 (2/3) (1/3) = 30 $(2/3)^2$ ~ 13.333333

E for aa = 30 $(1-\hat{p})^2$ = 30 $(1/3)^2$ ~ 3.333333

Estimating $\hat{p}$, necessary to reduce the expected entries to actual numbers, will cost us one degree of freedom. So the resulting chi-square will have df = 3 - 1 - 1 = 1.

	AA	Aa	aa
O	16	8	6
E	13.333333	13.333333	3.333333

The chi-square statistic works out to $\chi^2 = \sum_{\text{cells}} \frac{(O-E)^2}{E}$ = 4.8.

The P-value is (using a computer) for df = 1,

P-value = P(χ^2 > 4.8) ~ 0.0284597 (i.e. $t_{0.0284597}$ = 4.8).

It is therefore rather rare to encounter (as we have) a chi-square statistic with df = 1 as large or larger than 4.8. Either the model is incorrect or we have witnessed a rare event. Maybe not "bet your life on it" rare, but less than 3% rare.

Your table of chi-square has entries like $t_{.9}$ ~ 0.015791 and .$t_{0.1}$ ~ 2.70552 (see df = 1).