FORMULAS FOR REGRESSION AND HYPOTHESIS TESTING

\[ s_n^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 \]

\[ s_n^2 = \frac{n}{n-1} s_n^2 \]

\[ \hat{r} = \frac{\sum_{i=1}^{n} x_i y_i}{s_X s_Y} \]

\[ y = ax + b \text{ linear regression: } \hat{a} = \frac{\sum_{i=1}^{n} x_i y_i}{s_X^2}, \quad \hat{b} = \bar{y} - \hat{a} \bar{x} \]

**z-test:**

1. 1-sample, two-sided: \( z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \), \( z_{\alpha/2} = \Phi^{-1}(1 - \epsilon / 2) \), confidence interval for \( \mu \): \( \left[ \bar{x} - z_{\alpha/2} \frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2} \frac{\sigma}{\sqrt{n}} \right] \).
2. 1-sample, one-sided: \( z = \frac{\bar{x} - \mu}{\sigma / \sqrt{n}} \), \( z_{\epsilon} = \Phi^{-1}(1 - \epsilon) \).
3. 2-sample, two-sided: \( z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}} \), \( z_{\epsilon/2} = \Phi^{-1}(1 - \epsilon / 2) \).
4. 2-sample, one-sided: \( z = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{s_x^2}{n_1} + \frac{s_y^2}{n_2}}} \), \( z_{\epsilon} = \Phi^{-1}(1 - \epsilon) \).

**t-test:**

1. 1-sample, two-sided: \( t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \), \( t_{\epsilon/2} \) is the \( 1 - \epsilon / 2 \) quantile value of the \( t_{n-1} \)-distribution.
   confidence interval for \( \mu \): \( \left[ \bar{x} - t_{\epsilon/2} \frac{s}{\sqrt{n}}, \bar{x} + t_{\epsilon/2} \frac{s}{\sqrt{n}} \right] \).
2. 1-sample, one-sided: \( t = \frac{\bar{x} - \mu}{s / \sqrt{n}} \), \( t_{\epsilon} \) is the \( 1 - \epsilon \) quantile value of the \( t_{n-1} \)-distribution.
3. 2-sample, two-sided: \( t = \frac{\bar{x} - \bar{y}}{s_{X+Y} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sqrt{\frac{n_1 \nu_1}{n_1 + \nu_2} + \frac{n_2 \nu_2}{n_2 + \nu_1}}, \quad t_{\epsilon/2} \) is the \( 1 - \epsilon / 2 \) quantile value of the \( t_{n_1+n_2-2} \)-distribution.
4. 2-sample, one-sided: \( t = \frac{\bar{x} - \bar{y}}{s_{X+Y} \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sqrt{\frac{n_1 \nu_1}{n_1 + \nu_2} + \frac{n_2 \nu_2}{n_2 + \nu_1}}, \quad t_{\epsilon} \) is the \( 1 - \epsilon \) quantile value of the \( t_{n_1+n_2-2} \)-distribution.

**\( \chi^2 \)-test:**

1. Test for goodness of fit: \( \chi^2 = \sum_{i=1}^{r} \frac{(n_i - np_i)^2}{np_i} \) to be compared to the \( 1 - \epsilon \) quantile value of the \( \chi^2_{r-1} \)-distribution.
2. Test for homogeneity: \( \chi^2 = n \sum_{i=1}^{r} \frac{(\nu_i - np_i)^2}{\nu_i np_i} \) to be compared to the \( 1 - \epsilon \) quantile value of the \( \chi^2_{r-1} \)-distribution.
3. Test for independence: \( \chi^2 = \sum_{i=1}^{r} \sum_{j=1}^{r} \frac{(\nu_{i,j} - np_{i,j})^2}{np_{i,j}} \) to be compared to the \( 1 - \epsilon \) quantile value of the \( \chi^2_{(r-1)(r-1)} \)-distribution.

**Welch-test:**

\( t'(X, Y) = \sqrt{\frac{n_1 s_{X+Y}^2 / n_1 + n_2 s_{X+Y}^2 / n_2}{n_1 + n_2}} \), \( c = \frac{S_{X+Y}^2 / n_1 + S_{X+Y}^2 / n_2}{s_{X+Y}^2 / n_1 s_{X+Y}^2 / n_2}, \quad \frac{1 + \epsilon}{\gamma} = \frac{\epsilon^2}{n_1} + \frac{(1-\epsilon)^2}{n_2} \).

\( \chi_k^2 = \{ (x, y) \mid |t'(x, y)| \geq t_{\epsilon/2}(f) \} \) in the two-sided case,

\( \chi_k^2 = \{ (x, y) \mid t'(x, y) \geq t_\epsilon(f) \} \) in the one-sided case, where \( f \) is to be rounded to the nearest integer.