Statistical Inference on population means and variances

The information provided in this document are for reference only

3 conditions

A. Normal Distributions
B. Parameters Known
C. Large Sample Size

Legend for conditions: Y = Yes, N = No, -- = It does not matter

Confidence Interval

One Sample

Estimation of the population mean, μ

Conditions (A, B, C)	Statistic	Distribution	$100(1-\alpha)\%$ Confidence Interval
(Y, Y(σ^2), --)	$T = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	$N(0,1)$	$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$
(N, Y(σ^2), Y)	$T = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	Approximate $N(0,1)$ (CLT)	$\bar{X} \pm z_{\alpha/2}\frac{\sigma}{\sqrt{n}}$
(Y, N, --)	$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$	$t(n-1)$	$\bar{X} \pm t_{n-1;\alpha/2}\frac{S}{\sqrt{n}}$
(N, N, Y)	$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$	Approximate $N(0,1)$ (CLT & LLN)	$\bar{X} \pm z_{\alpha/2}\frac{S}{\sqrt{n}}$

* Y = Yes, N = No, - = It does not matter

CLT: Central Limit Theorem, LLN = Law of Large Numbers

$\Pr(Z > z_{\alpha/2}) = \alpha/2$ with $Z \sim N(0,1)$, $\Pr(T > t_{v;\,\alpha/2}) = \alpha/2$ with $T \sim t(v)$

Estimation of the population variance, σ^2

Conditions (A, B, C)	Statistic	Distribution	$100(1-\alpha)\%$ Confidence Interval
(Y, Y(μ), --)	$T = \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\sigma^2}$	$\chi^2(n)$	$\frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{n;\,\alpha/2}} < \sigma^2 < \frac{\sum_{i=1}^{n}(X_i-\mu)^2}{\chi^2_{n;1-\alpha/2}}$
(Y, N, N)	$T = \frac{(n-1)S^2}{\sigma^2}$	$\chi^2(n-1)$	$\frac{(n-1)S^2}{\chi^2_{n-1;\,\alpha/2}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{n-1;1-\alpha/2}}$

$\Pr(W > \chi^2_{(v;\,\alpha)}) = \alpha$ with $W \sim \chi^2(v)$

Two Samples

Paired Samples

Consider $D_i = X_i - Y_i$. Let $\mu_D = \mu_X - \mu_Y$

Estimation of the population mean, μ_D

Conditions (A, B, C)	Statistic	Distribution	$100(1-\alpha)\%$ Confidence Interval
(Y, N, N)	$T = \dfrac{\bar{X}_D - \mu_D}{S_D/\sqrt{n}}$	$t(n-1)$	$\bar{X}_D \pm t_{n-1;\alpha/2} \dfrac{S_D}{\sqrt{n}}$
(N, N, Y)	$T = \dfrac{\bar{X}_D - \mu_D}{S_D/\sqrt{n}}$	Approximate $N(0,1)$ (CLT & LLN)	$\bar{X}_D \pm z_{\alpha/2} \dfrac{S_D}{\sqrt{n}}$

Two Independent Samples

Estimation of the difference of population means, $\mu_1 - \mu_2$

Conditions (A, B, C)	Statistic	Distribution	$100(1-\alpha)\%$ Confidence Interval
(Y, Y(σ_1^2, σ_2^2), --)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$	$N(0,1)$	$\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$
(N, Y(σ_1^2, σ_2^2), Y)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$	Approximate $N(0,1)$	$\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2}\sqrt{\dfrac{\sigma_1^2}{n_1} + \dfrac{\sigma_2^2}{n_2}}$
(Y, N($\sigma_1^2 = \sigma_2^2$), --)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{S_p^2(1/n_1 + 1/n_2)}}$	$t(n_1 + n_2 - 2)$	$\bar{X}_1 - \bar{X}_2 \pm t_{n_1+n_2-2;\alpha/2}\sqrt{S_p^2\left(\dfrac{1}{n_1} + \dfrac{1}{n_2}\right)}$
(N, N, Y)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - (\mu_1 - \mu_2)}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$	Approximate $N(0,1)$ (CLT & LLN)	$\bar{X}_1 - \bar{X}_2 \pm z_{\alpha/2}\sqrt{\dfrac{S_1^2}{n_1} + \dfrac{S_2^2}{n_2}}$

Estimation of the ratio of population variances, σ_1^2/σ_2^2

Conditions (A, B, C)	Statistic	Distribution	$100(1-\alpha)\%$ Confidence Interval
(Y, N, --)	$T = \dfrac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2}$	F_{n_1-1,n_2-1}	$\dfrac{S_1^2}{S_2^2}\dfrac{1}{F_{n_1-1,n_2-1;\,\alpha/2}} < \dfrac{\sigma_1^2}{\sigma_2^2} < \dfrac{S_1^2}{S_2^2}F_{n_2-1,n_1-1;\,\alpha/2}$ Or $\dfrac{S_1^2}{S_2^2}\dfrac{1}{F_{n_1-1,n_2-1;\,\alpha/2}} < \dfrac{\sigma_1^2}{\sigma_2^2} < \dfrac{S_1^2}{S_2^2}\dfrac{1}{F_{n_1-1,n_2-1;\,1-\alpha/2}}$

$\Pr\left(W > F_{v_1,v_2;\,\alpha}\right) = \alpha$ with $W \sim F(v_1, v_2)$

Hypothesis Testing
One Sample

Hypothesis testing on the population mean. H0: $\mu = \mu_0$

Conditions (A, B, C)	Test Statistic	Alternative	Reject H0 if
(Y, Y(σ^2), --)	$T = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	H1: $\mu \neq \mu_0$	$T < -z_{\alpha/2}$ or $T > z_{\alpha/2}$
		H1: $\mu > \mu_0$	$T > z_{\alpha}$
		H1: $\mu < \mu_0$	$T < -z_{\alpha}$
(N, Y(σ^2), Y)	$T = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}}$	H1: $\mu \neq \mu_0$	$T < -z_{\alpha/2}$ or $T > z_{\alpha/2}$
		H1: $\mu > \mu_0$	$T > z_{\alpha}$
		H1: $\mu < \mu_0$	$T < -z_{\alpha}$
(Y, N, --)	$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	H1: $\mu \neq \mu_0$	$T < -t_{n-1;\,\alpha/2}$ or $T > t_{n-1;\,\alpha/2}$
		H1: $\mu > \mu_0$	$T > t_{n-1;\,\alpha}$
		H1: $\mu < \mu_0$	$T < -t_{n-1;\,\alpha}$
(N, N, Y)	$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}}$	H1: $\mu \neq \mu_0$	$T < -z_{\alpha/2}$ or $T > z_{\alpha/2}$
		H1: $\mu > \mu_0$	$T > z_{\alpha}$
		H1: $\mu < \mu_0$	$T < -z_{\alpha}$

* Y = Yes, N = No, - = It does not matter

Hypothesis testing on the population variance. H0: $\sigma^2 = \sigma_0^2$

Conditions (A, B, C)	Test Statistic	Alternative	Reject H0 if
(Y, N, --)	$T = \frac{(n-1)S^2}{\sigma_0^2}$	H1: $\sigma^2 \neq \sigma_0^2$	$T < \chi^2_{(n-1;1-\alpha/2)}$ or $T > \chi^2_{(n-1;\,\alpha/2)}$
		H1: $\sigma^2 > \sigma_0^2$	$T > \chi^2_{(n-1;\,\alpha)}$
		H1: $\sigma^2 < \sigma_0^2$	$T < \chi^2_{(n-1;\,1-\alpha)}$

$\Pr\left(W > \chi^2_{(v;\,\alpha)}\right) = \alpha$ with $W \sim \chi^2(v)$

Two Samples

Paired Samples

Consider $D_i = X_i - Y_i$. Let $\mu_D = \mu_X - \mu_Y$

Hypothesis testing on the population mean, H0: $\mu_D = 0$

Conditions (A, B, C)	Test Statistic	Alternative	Reject H0 if
(Y, N, --)	$T = \frac{\bar{X}_D - 0}{S_D/\sqrt{n}}$	H1: $\mu_D \neq 0$	$T < -t_{n-1;\,\alpha/2}$ or $T > t_{n-1;\,\alpha/2}$
		H1: $\mu_D > 0$	$T > t_{n-1;\,\alpha}$
		H1: $\mu_D < 0$	$T < -t_{n-1;\,\alpha}$
(N, N, Y)	$T = \frac{\bar{X}_D - 0}{S_D/\sqrt{n}}$	H1: $\mu_D \neq 0$	$T < -z_{\alpha/2}$ or $T > z_{\alpha/2}$
		H1: $\mu_D > 0$	$T > z_{\alpha}$
		H1: $\mu_D < 0$	$T < -z_{\alpha}$

Two Independent Samples

Hypothesis testing on the difference of population means, H0: $\mu_1 - \mu_2 = 0$

Conditions (A, B, C)	Test Statistic	Alternative	Reject H0 if
(Y, Y(σ_1^2, σ_2^2), --)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$	H1: $\mu_1 - \mu_2 \neq 0$	$T < -z_{\alpha/2}$ or $T > z_{\alpha/2}$
		H1: $\mu_1 - \mu_2 > 0$	$T > z_{\alpha}$
		H1: $\mu_1 - \mu_2 < 0$	$T < -z_{\alpha}$
(N, Y(σ_1^2, σ_2^2), Y)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{\sigma_1^2/n_1 + \sigma_2^2/n_2}}$	H1: $\mu_1 - \mu_2 \neq 0$	$T < -z_{\alpha/2}$ or $T > z_{\alpha/2}$
		H1: $\mu_1 - \mu_2 > 0$	$T > z_{\alpha}$
		H1: $\mu_1 - \mu_2 < 0$	$T < -z_{\alpha}$
(Y, N($\sigma_1^2 = \sigma_2^2$), --)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{S_p^2(1/n_1 + 1/n_2)}}$	H1: $\mu_1 - \mu_2 \neq 0$	$T < -t_{n_1+n_2-2;\, \alpha/2}$ or $T > t_{n_1+n_2-2;\, \alpha/2}$
		H1: $\mu_1 - \mu_2 > 0$	$T > t_{n_1+n_2-2;\, \alpha}$
		H1: $\mu_1 - \mu_2 < 0$	$T < -t_{n_1+n_2-2;\, \alpha}$
(N, N, Y)	$T = \dfrac{\bar{X}_1 - \bar{X}_2 - 0}{\sqrt{S_1^2/n_1 + S_2^2/n_2}}$	H1: $\mu_1 - \mu_2 \neq 0$	$T < -z_{\alpha/2}$ or $T > z_{\alpha/2}$
		H1: $\mu_1 - \mu_2 > 0$	$T > z_{\alpha}$
		H1: $\mu_1 - \mu_2 < 0$	$T < -z_{\alpha}$

Hypothesis testing on the ratio of population variances, H0: $\sigma_1^2/\sigma_2^2 = 1$

Conditions (A, B, C)	Test Statistic	Alternative	Reject H0 if
(Y, N, --)	$T = \dfrac{S_1^2}{S_2^2}$	H1: $\sigma_1^2/\sigma_2^2 \neq 1$	$F < F_{n_1-1,n_2-1;\, 1-\alpha/2}$ or $F > F_{n_1-1,n_2-1;\, \alpha/2}$
		H1: $\sigma_1^2/\sigma_2^2 > 1$	$F > F_{n_1-1,n_2-1;\alpha}$
		H1: $\sigma_1^2/\sigma_2^2 < 1$	$F < F_{n_1-1,n_2-1;1-\alpha}$ $(= 1/F_{n_2-1,n_1-1;\alpha})$

$\Pr\left(W > F_{v_1,v_2;\,\alpha}\right) = \alpha$ with $W \sim F(v_1, v_2)$