18.650
Statistics for Applications

Chapter 5: Parametric hypothesis testing

Cherry Blossom run (1)

- The credit union Cherry Blossom Run is a 10 mile race that takes place every year in D.C.
- In 2009 there were 14974 participants
- Average running time was 103.5 minutes.

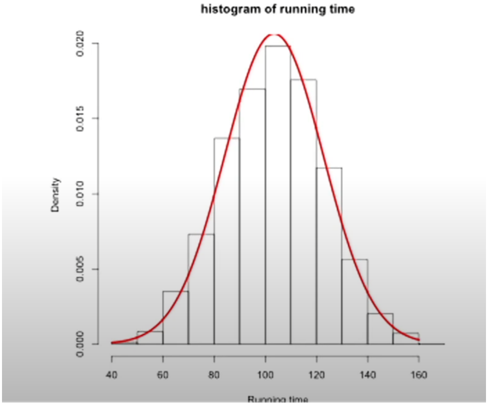
Were runners faster in 2012?

To answer this question, select n runners from the 2012 race at random and denote by $X_1, \ldots, X_n$ their running time.

Cherry Blossom run (2)

We can see from past data that the running time has Gaussian distribution.

The variance was 373.



Cherry Blossom run (3)

- We are given i.i.d r.v $X_1, \ldots, X_n$ and we want to know if $X_1 \sim \mathcal{N}(103.5, 373)$
- This is a **hypothesis testing** problem.
- There are many ways this could be false:
  1. $\mathbb{E}[X_1] \neq 103.5$
  2. $\mathsf{var}[X_1] \neq 373$
  3. X_1 may not even be Gaussian.
- We are interested in a very specific question: is $\mathbb{E}[X_1] < 103.5$?

Cherry Blossom run (4)

- We make the following **assumptions**:
  1. var$[X_1] = 373$ (variance is the same between 2009 and 2012)
  2. X_1 is Gaussian.
- The only thing that we did not fix is $\mathbb{E}[X_1] = \mu$.
- Now we want to test (only): "Is $\mu = 103.5$ or is $\mu < 103.5$"?
- By making **modeling assumptions**, we have reduced the number of ways the hypothesis $X_1 \sim \mathcal{N}(103.5, 373)$ may be rejected.
- The only way it can be rejected is if $X_1 \sim \mathcal{N}(\mu, 373)$ for some $\mu < 103.5$.
- We compare an expected value to a fixed reference number (103.5).

Cherry Blossom run (5)

Simple heuristic:

$$\text{“If } \bar{X}_n < 103.5\text{, then } \mu < 103.5\text{”}$$

This could go wrong if I randomly pick only fast runners in my sample $X_1, \ldots, X_n$.

Better heuristic:

$$\text{“If } \bar{X}_n < 103.5 - (\text{something that } \xrightarrow[n\to\infty]{} 0)\text{, then } \mu < 103.5\text{”}$$

To make this intuition more precise, we need to take the size of the random fluctuations of $\bar{X}_n$ into account!

Clinical trials (1)

- Pharmaceutical companies use hypothesis testing to test if a new drug is efficient.
- To do so, they administer a drug to a group of patients (test group) and a placebo to another group (control group).
- Assume that the drug is a cough syrup.
- Let μ_{control} denote the expected number of expectorations per hour after a patient has used the placebo.
- Let μ_{drug} denote the expected number of expectorations per hour after a patient has used the syrup.
- We want to know if $\mu_{\text{drug}} < \mu_{\text{control}}$
- We compare two expected values. No reference number.

Clinical trials (2)

- Let $X_1, \ldots, X_{n_{\mathrm{drug}}}$ denote n_{drug} i.i.d r.v. with distribution Poiss(μ_{drug})
- Let $Y_1, \ldots, Y_{n_{\mathrm{control}}}$ denote n_{control} i.i.d r.v. with distribution Poiss(μ_{control})
- We want to test if $\mu_{\mathrm{drug}} < \mu_{\mathrm{control}}$.

Heuristic:

"If $\bar{X}_{\mathrm{drug}} < \bar{X}_{\mathrm{control}} -$(something that $\xrightarrow[\substack{n_{\mathrm{drug}} \to \infty \\ n_{\mathrm{control}} \to \infty}]{} 0$), then conclude that $\mu_{\mathrm{drug}} < \mu_{\mathrm{control}}$"

Heuristics (1)

Example 1: A coin is tossed 80 times, and Heads are obtained 54 times. Can we conclude that the coin is significantly unfair ?

- $n = 80$, $X_1, \ldots, X_n \overset{iid}{\sim} \mathsf{Ber}(p)$;
- $\bar{X}_n = 54/80 = .68$
- If it was true that $p = .5$: By CLT+Slutsky's theorem,

$$\sqrt{n}\frac{\bar{X}_n - .5}{\sqrt{.5(1-.5)}} \approx \mathcal{N}(0,1).$$

- $\sqrt{n}\frac{\bar{X}_n - .5}{\sqrt{.5(1-.5)}} \approx 3.22$
- Conclusion: It **seems quite** reasonable to reject the hypothesis $p = .5$.

Heuristics (2)

Example 2: A coin is tossed 30 times, and Heads are obtained 13 times. Can we conclude that the coin is significantly unfair ?

- $n = 30, X_1, \ldots, X_n \overset{iid}{\sim} \mathsf{Ber}(p)$;
- $\bar{X}_n = 13/30 \approx .43$
- If it was true that $p = .5$: By CLT+Slutsky's theorem,

$$\sqrt{n}\frac{\bar{X}_n - .5}{\sqrt{.5(1-.5)}} \approx \mathcal{N}(0,1).$$

- Our data gives $\sqrt{n}\frac{\bar{X}_n - .5}{\sqrt{.5(1-.5)}} \approx -.77$
- The number $.77$ is a plausible realization of a random variable $Z \sim \mathcal{N}(0,1)$.
- Conclusion: our data does not suggest that the coin is unfair.

Statistical formulation (1)

- Consider a sample $X_1, \ldots, X_n$ of i.i.d. random variables and a statistical model $(E, (\mathbb{P}_\theta)_{\theta \in \Theta})$.
- Let Θ_0 and Θ_1 be disjoint subsets of Θ.
- Consider the two hypotheses: $\begin{cases} H_0 : & \theta \in \Theta_0 \\ H_1 : & \theta \in \Theta_1 \end{cases}$
- H_0 is the *null hypothesis*, H_1 is the *alternative hypothesis*.
- If we believe that the true θ is either in Θ_0 or in Θ_1, we may want to *test* H_0 *against* H_1.
- We want to decide whether to *reject* H_0 (look for evidence against H_0 in the data).

Statistical formulation (2)

- H_0 and H_1 do not play a symmetric role: the data is is only used to try to disprove H_0
- In particular lack of evidence, does not mean that H_0 is true ("innocent until proven guilty")
- A *test* is a statistic $\psi \in \{0, 1\}$ such that:
  - If $\psi = 0$, H_0 is not rejected;
  - If $\psi = 1$, H_0 is rejected.
- Coin example: H_0: $p = 1/2$ vs. H_1: $p = 1/2$.
- $\psi = \mathbb{1}\left\{ \left| \sqrt{n} \frac{\bar{X}_n - .5}{\sqrt{.5(1 - .5)}} \right| > C \right\}$, for some $C > 0$.
- How to choose the *threshold* C ?

Statistical formulation (3)

- *Rejection region* of a test ψ:

$$R_\psi = \{x \in E^n : \psi(x) = 1\}.$$

- *Type 1 error* of a test ψ (rejecting H_0 when it is actually true):

$$\begin{array}{cccc} \alpha_\psi & : & \Theta_0 & \to & \mathbb{R} \\ & & \theta & \mapsto & \mathbb{P}_\theta[\psi = 1]. \end{array}$$

- *Type 2 error* of a test ψ (not rejecting H_0 although H_1 is actually true):

$$\begin{array}{cccc} \beta_\psi & : & \Theta_1 & \to & \mathbb{R} \\ & & \theta & \mapsto & \mathbb{P}_\theta[\psi = 0]. \end{array}$$

- *Power* of a test ψ:

$$\pi_\psi = \inf_{\theta \in \Theta_1} \left(1 - \beta_\psi(\theta)\right).$$

Statistical formulation (4)

- A test ψ has *level* α if
$$\alpha_\psi(\theta) \le \alpha, \qquad \forall \theta \in \Theta_0.$$
- A test ψ has *asymptotic level* α if
$$\lim_{n\to\infty} \alpha_\psi(\theta) \le \alpha, \qquad \forall \theta \in \Theta_0.$$
- In general, a test has the form
$$\psi = \mathbb{1}\{T_n > c\},$$
for some statistic T_n and threshold $c \in \mathbb{R}$.
- T_n is called the *test statistic*. The rejection region is $R_\psi = \{T_n > c\}$.

Example (1)

- Let $X_1, \ldots, X_n \overset{iid}{\sim} \mathsf{Ber}(p)$, for some unknown $p \in (0,1)$.
- We want to test:

$$H_0\colon\ p = 1/2 \text{ vs. } H_1\colon\ p = 1/2$$

  with asymptotic level $\alpha \in (0,1)$.

- Let $T_n = \sqrt{n}\ \dfrac{\hat{p}_n - 0.5}{\sqrt{.5(1-.5)}}$, where $\hat{p}_n$ is the MLE.

- If H_0 is true, then by CLT and Slutsky's theorem,

$$\mathbb{P}[T_n > q_{\alpha/2}] \xrightarrow[n\to\infty]{} 0.05$$

- Let $\psi_\alpha = \mathbb{1}\{T_n > q_{\alpha/2}\}$.

Example (2)

Coming back to the two previous coin examples: For $\alpha = 5\%$, $q_{\alpha/2} = 1.96$, so:

- In **Example 1**, H_0 is rejected at the asymptotic level 5% by the test $\psi_{5\%}$;
- In **Example 2**, H_0 is not rejected at the asymptotic level 5% by the test $\psi_{5\%}$.

Question: In **Example 1**, for what level α would ψ_α not reject H_0 ? And in **Example 2**, at which level α would ψ_α reject H_0 ?

p-value

Definition

The (asymptotic) *p-value* of a test ψ_α is the smallest (asymptotic) level α at which ψ_α rejects H_0. It is random, it depends on the sample.

Golden rule

p-value $\leq \alpha \quad \Leftrightarrow \quad H_0$ is rejected by ψ_α, at the (asymptotic) level α.

The smaller the p-value, the more confidently one can reject H_0.

- Example 1: p-value $= \mathbb{P}[|Z| > 3.21] \ll .01$.
- Example 2: p-value $= \mathbb{P}[|Z| > .77] \approx .44$.

Neyman-Pearson's paradigm

Idea: For given hypotheses, among all tests of level/asymptotic level α, is it possible to find one that has maximal power ?

Example: The trivial test $\psi = 0$ that never rejects H_0 has a perfect level ($\alpha = 0$) but poor power ($\pi_\psi = 0$).

Neyman-Pearson's theory provides (the most) powerful tests with given level. In 18.650, we only study several cases.

The χ^2 distributions

Definition

For a positive integer d, the χ^2 *(pronounced "Kai-squared") distribution with d degrees of freedom* is the law of the random variable $Z_1^2 + Z_2^2 + \ldots + Z_d^2$, where $Z_1, \ldots, Z_d \overset{iid}{\sim} \mathcal{N}(0,1)$.

Examples:

- If $Z \sim \mathcal{N}_d(\mathbf{0}, I_d)$, then $\|Z\|_2^2 \sim \chi_d^2$.
- Recall that the sample variance is given by
  $S_n = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 - (\bar{X}_n)^2$
- Cochran's theorem implies that for $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, if S_n is the sample variance, then
  $$\frac{nS_n}{\sigma^2} \sim \chi_{n-1}^2.$$
- $\chi_2^2 = \mathsf{Exp}(1/2)$.

Student's T distributions

Definition

For a positive integer d, the *Student's T distribution with d degrees of freedom* (denoted by t_d) is the law of the random variable $\dfrac{Z}{\sqrt{V/d}}$, where $Z \sim \mathcal{N}(0,1)$, $V \sim \chi^2_d$ and $Z \perp\!\!\!\perp V$ (Z is independent of V).

Example:

- Cochran's theorem implies that for $X_1, \ldots, X_n \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, if S_n is the sample variance, then

$$\sqrt{n-1}\ \frac{\bar{X}_n - \mu}{\sqrt{S_n}} \sim t_{n-1}.$$

Wald's test (1)

- Consider an i.i.d. sample $X_1, \ldots, X_n$ with statistical model $(E, (\mathbb{P}_\theta)_{\theta \in \Theta})$, where $\Theta \subseteq \mathbb{R}^d$ $(d \geq 1)$ and let $\theta_0 \in \Theta$ be fixed and given.
- Consider the following hypotheses:

$$\begin{cases} H_0 : & \theta = \theta_0 \\ H_1 : & \theta = \theta_0. \end{cases}$$

- Let $\hat{\theta}^{MLE}$ be the MLE. Assume the MLE technical conditions are satisfied.

- If H_0 is true, then

$$\sqrt{n}\ I(\hat{\theta}^{MLE})^{1/2} \left(\hat{\theta}_n^{MLE} - \theta_0 \right) \xrightarrow[n \to \infty]{(d)} \mathcal{N}_d\left(0, I_d\right) \qquad \text{w.r.t. } \mathbb{P}_{\theta_0}.$$

Wald's test (2)

- Hence,

$$\underbrace{n\left(\hat{\theta}_n^{MLE} - \theta_0\right)^\top I(\hat{\theta}^{MLE})\left(\hat{\theta}_n^{MLE} - \theta_0\right)}_{T_n} \xrightarrow[n\to\infty]{(d)} \chi^2_d \quad \text{w.r.t. } \mathbb{P}_{\theta_0}.$$

- Wald's test with asymptotic level $\alpha \in (0,1)$:

$$\psi = \mathbb{1}\{T_n > q_\alpha\},$$

  where q_α is the $(1-\alpha)$-quantile of χ^2_d (see tables).

- Remark: Wald's test is also valid if H_1 has the form "$\theta > \theta_0$" or "$\theta < \theta_0$" or "$\theta = \theta_1$"...

Likelihood ratio test (1)

- Consider an i.i.d. sample $X_1, \ldots, X_n$ with statistical model $(E, (\mathbb{P}_\theta)_{\theta \in \Theta})$, where $\Theta \subseteq \mathbb{R}^d$ $(d \geq 1)$.

- Suppose the null hypothesis has the form

$$H_0 : (\theta_{r+1}, \ldots, \theta_d) = (\theta_{r+1}^{(0)}, \ldots, \theta_d^{(0)}),$$

  for some fixed and given numbers $\theta_{r+1}^{(0)}, \ldots, \theta_d^{(0)}$.

- Let

$$\hat{\theta}_n = \underset{\theta \in \Theta}{\operatorname{argmax}} \ \ell_n(\theta) \quad \text{(MLE)}$$

  and

$$\hat{\theta}_n^c = \underset{\theta \in \Theta_0}{\operatorname{argmax}} \ \ell_n(\theta) \quad \text{(``constrained MLE'')}$$

Likelihood ratio test (2)

- Test statistic:

$$T_n = 2\left(\ell_n(\hat{\theta}_n) - \ell_n(\hat{\theta}_n^c)\right).$$

- Theorem

  Assume H_0 is true and the MLE technical conditions are satisfied. Then,

$$T_n \xrightarrow[n\to\infty]{(d)} \chi^2_{d-r} \quad \text{w.r.t. } \mathbb{P}_\theta.$$

- Likelihood ratio test with asymptotic level $\alpha \in (0,1)$:

$$\psi = \mathbb{1}\{T_n > q_\alpha\},$$

  where q_α is the $(1-\alpha)$-quantile of χ^2_{d-r} (see tables).

Testing implicit hypotheses (1)

- Let $X_1, \ldots, X_n$ be i.i.d. random variables and let $\theta \in \mathbb{R}^d$ be a parameter associated with the distribution of X_1 (e.g. a moment, the parameter of a statistical model, etc...)

- Let $g : \mathbb{R}^d \to \mathbb{R}^k$ be continuously differentiable (with $k < d$).

- Consider the following hypotheses:

$$\begin{cases} H_0 : & g(\theta) = 0 \\ H_1 : & g(\theta) = 0. \end{cases}$$

- E.g. $g(\theta) = (\theta_1, \theta_2)$ $(k = 2)$, or $g(\theta) = \theta_1 - \theta_2$ $(k = 1)$, or...

Testing implicit hypotheses (2)

- Suppose an asymptotically normal estimator $\hat{\theta}_n$ is available:

$$\sqrt{n} \quad \hat{\theta}_n - \theta \quad \xrightarrow[n\to\infty]{(d)} \mathcal{N}_d(0, \Sigma(\theta)).$$

- Delta method:

$$\sqrt{n} \quad g(\hat{\theta}_n) - g(\theta) \quad \xrightarrow[n\to\infty]{(d)} \mathcal{N}_k\left(0, \Gamma(\theta)\right),$$

where $\Gamma(\theta) = \nabla g(\theta)^\top \Sigma(\theta) \nabla g(\theta) \in \mathbb{R}^{k\times k}$.

- Assume $\Sigma(\theta)$ is invertible and $\nabla g(\theta)$ has rank k. So, $\Gamma(\theta)$ is invertible and

$$\sqrt{n}\, \Gamma(\theta)^{-1/2} \quad g(\hat{\theta}_n) - g(\theta) \quad \xrightarrow[n\to\infty]{(d)} \mathcal{N}_k\left(0, I_k\right).$$

Testing implicit hypotheses (3)

- Then, by Slutsky's theorem, if $\Gamma(\theta)$ is continuous in θ,

$$\sqrt{n}\,\Gamma(\hat{\theta}_n)^{-1/2} \quad g(\hat{\theta}_n) - g(\theta) \quad \xrightarrow[n\to\infty]{(d)} \mathcal{N}_k\left(0, I_k\right).$$

- Hence, if H_0 is true, i.e., $g(\theta) = 0$,

$$\underbrace{n g(\hat{\theta}_n)^\top \Gamma^{-1}(\hat{\theta}_n) g(\hat{\theta}_n)}_{T_n} \xrightarrow[n\to\infty]{(d)} \chi^2_k.$$

- Test with asymptotic level α:

$$\psi = \mathbb{1}\{T_n > q_\alpha\},$$

where q_α is the $(1-\alpha)$-quantile of χ^2_k (see tables).

The multinomial case: χ^2 test (1)

Let $E = \{a_1, \dots, a_K\}$ be a finite space and $(\mathbb{P}_{\mathbf{p}})_{\mathbf{p} \in \Delta_K}$ be the family of all probability distributions on E:

- $$\Delta_K = \left\{ \mathbf{p} = (p_1, \dots, p_K) \in (0,1)^K : \sum_{j=1}^{K} p_j = 1 \right\}.$$
- For $\mathbf{p} \in \Delta_K$ and $X \sim \mathbb{P}_{\mathbf{p}}$,

$$\mathbb{P}_{\mathbf{p}}[X = a_j] = p_j, \quad j = 1, \dots, K.$$

The multinomial case: χ^2 test (2)

- Let $X_1, \ldots, X_n \overset{iid}{\sim} \mathbb{P}_{\mathbf{p}}$, for some unknown $\mathbf{p} \in \Delta_K$, and let $\mathbf{p}^0 \in \Delta_K$ be fixed.
- We want to test:
$$H_0\colon\ \mathbf{p} = \mathbf{p}^0 \text{ vs. } H_1\colon\ \mathbf{p} = \mathbf{p}^0$$
with asymptotic level $\alpha \in (0, 1)$.
- Example: If $\mathbf{p}^0 = (1/K, 1/K, \ldots, 1/K)$, we are testing whether $\mathbb{P}_{\mathbf{p}}$ is the uniform distribution on E.

The multinomial case: χ^2 test (3)

- Likelihood of the model:

$$L_n(X_1, \ldots, X_n, \mathbf{p}) = p_1^{N_1} p_2^{N_2} \ldots p_K^{N_K},$$

where $N_j = \#\{i = 1, \ldots, n : X_i = a_j\}$.

- Let $\hat{\mathbf{p}}$ be the MLE:

$$\hat{\mathbf{p}}_j = \frac{N_j}{n}, \qquad j = 1, \ldots, K.$$

⚠ $\hat{\mathbf{p}}$ maximizes $\log L_n(X_1, \ldots, X_n, \mathbf{p})$ **under the constraint**

$$\sum_{j=1}^{K} p_j = 1.$$

The multinomial case: χ^2 test (4)

- If H_0 is true, then $\sqrt{n}(\hat{\mathbf{p}} - \mathbf{p}^0)$ is asymptotically normal, and the following holds.

Theorem

$$n \underbrace{\sum_{j=1}^{K} \frac{\left(\hat{\mathbf{p}}_j - \mathbf{p}_j^0\right)^2}{\mathbf{p}_j^0}}_{T_n} \xrightarrow[n\to\infty]{(d)} \chi^2_{K-1}.$$

- χ^2 test with asymptotic level α: $\quad \psi_\alpha = \mathbb{1}\{T_n > q_\alpha\}$, where q_α is the $(1-\alpha)$-quantile of χ^2_{K-1}.
- Asymptotic p-value of this test: $p - \text{value} = \mathbb{P}\left[Z > T_n | T_n\right]$, where $Z \sim \chi^2_{K-1}$ and $Z \perp\!\!\!\perp T_n$.

The Gaussian case: Student's test (1)

- Let $X_1, \ldots, X_n \stackrel{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$, for some unknown $\mu \in \mathbb{R}, \sigma^2 > 0$ and let $\mu_0 \in \mathbb{R}$ be fixed, given.

- We want to test:

$$H_0\colon\ \mu = \mu_0 \text{ vs. } H_1\colon\ \mu = \mu_0$$

  with asymptotic level $\alpha \in (0, 1)$.

- **If σ^2 is known:** Let $T_n = \sqrt{n}\ \dfrac{\bar{X}_n - \mu_0}{\sigma}$. Then, $T_n \sim \mathcal{N}(0, 1)$ and

$$\psi_\alpha = \mathbb{1}\{|T_n| > q_{\alpha/2}\}$$

  is a test with (non asymptotic) level α.

The Gaussian case: Student's test (2)

If σ^2 is unknown:

- Let $\widetilde{T_n} = \sqrt{n-1}\ \dfrac{\bar{X}_n - \mu_0}{\sqrt{S_n}}$, where S_n is the sample variance.

- Cochran's theorem:
  - $\bar{X}_n \perp\!\!\!\perp S_n$;
  - $\dfrac{nS_n}{\sigma^2} \sim \chi^2_{n-1}$.

- Hence, $\widetilde{T_n} \sim t_{n-1}$: Student's distribution with $n-1$ degrees of freedom.

The Gaussian case: Student's test (3)

- Student's test with (non asymptotic) level $\alpha \in (0,1)$:

$$\psi_\alpha = \mathbb{1}\{|\widetilde{T_n}| > q_{\alpha/2}\},$$

  where $q_{\alpha/2}$ is the $(1-\alpha/2)$-quantile of t_{n-1}.

- If H_1 is $\mu > \mu_0$, Student's test with level $\alpha \in (0,1)$ is:

$$\psi'_\alpha = \mathbb{1}\{\widetilde{T_n} > q_\alpha\},$$

  where q_α is the $(1-\alpha)$-quantile of t_{n-1}.

- Advantage of Student's test:
  - Non asymptotic
  - Can be run on small samples

- Drawback of Student's test: It relies on the assumption that the sample is Gaussian.

Two-sample test: large sample case (1)

- Consider two samples: $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_m$, of independent random variables such that

$$\mathbb{E}[X_1] = \cdots = \mathbb{E}[X_n] = \mu_X$$

, and

$$\mathbb{E}[Y_1] = \cdots = \mathbb{E}[Y_m] = \mu_Y$$

- Assume that the variances of are known so assume (without loss of generality) that

$$\mathsf{var}(X_1) = \cdots = \mathsf{var}(X_n) = \mathsf{var}(Y_1) = \cdots = \mathsf{var}(Y_m) = 1$$

- We want to test:

$$H_0\text{: } \mu_X = \mu_Y \text{ vs. } H_1\text{: } \mu_X = \mu_Y$$

with asymptotic level $\alpha \in (0, 1)$.

Two-sample test: large sample case (2)

From CLT:

$$\sqrt{n}(\bar{X}_n - \mu_X) \xrightarrow[n\to\infty]{(d)} \mathcal{N}(0,1)$$

and

$$\sqrt{m}(\bar{Y}_m - \mu_Y) \xrightarrow[m\to\infty]{(d)} \mathcal{N}(0,1) \quad \Rightarrow \quad \sqrt{n}(\bar{Y}_m - \mu_Y) \xrightarrow[\substack{n\to\infty \\ m\to\infty \\ \frac{m}{n}\to\gamma}]{(d)} \mathcal{N}(0,\gamma)$$

Moreover, the two samples are independent so

$$\sqrt{n}(\bar{X}_n - \bar{Y}_m) + \sqrt{n}(\mu_X - \mu_Y) \xrightarrow[\substack{n\to\infty \\ m\to\infty \\ \frac{m}{n}\to\gamma}]{(d)} \mathcal{N}(0, 1+\gamma)$$

Under $H_0 : \mu_X = \mu_Y$:

$$\sqrt{n}\frac{\bar{X}_n - \bar{Y}_m}{\sqrt{1+m/n}} \xrightarrow[\substack{n\to\infty \\ m\to\infty \\ \frac{m}{n}\to\gamma}]{(d)} \mathcal{N}(0,1)$$

Test: $$\psi_\alpha = \mathbb{1}\left\{ \sqrt{n}\frac{\bar{X}_n - \bar{Y}_m}{\sqrt{1+m/n}} > q_{\alpha/2}\right\}$$

Two-sample T-test

- If the variances are unknown but we know that $X_i \sim \mathcal{N}(\mu_X, \sigma_X^2),\ Y_i \sim \mathcal{N}(\mu_Y, \sigma_Y^2)$.
- Then
$$\bar{X}_n - \bar{Y}_m \sim \mathcal{N}\big(\mu_X - \mu_Y, \frac{\sigma_X^2}{n} + \frac{\sigma_Y^2}{m}\big)$$
- Under H_0:
$$\frac{\bar{X}_n - \bar{Y}_m}{\sqrt{\sigma_X^2/n + \sigma_Y^2/m}} \sim \mathcal{N}(0,1)$$
- For unknown variance:
$$\frac{\bar{X}_n - \bar{Y}_m}{\sqrt{S_X^2/n + S_Y^2/m}} \sim t_N$$

  where
$$N = \frac{\left(S_X^2/n + S_Y^2/m\right)^2}{\frac{S_X^4}{n^2(n-1)} + \frac{S_Y^4}{m^2(m-1)}}$$