Stat 205: Introduction to Nonparametric Statistics

Lecture 02: Nonparametric Inference

Instructor David Donoho; TA: Yu Wang

What is Nonparametric?, 1

Nonparametric = nonParametric.

- Not $N(\mu, \sigma^2)$
- Not $Y = \alpha + \beta X + Z$

Two meanings, two very different use cases.

Inference techniques offering valid p-values and confidence statements under minimal assumptions: sign test, median test, Wilcoxon's tests, and the Kruskal-Wallis and Friedman tests, tests of independence.

Predictive modeling techniques valid quite generally, such as kernel and spline smoothing, nearest neighbor, and even deep neural networks.

What is Nonparametric?, 2

Nonparametric = nonParametric.

- Not $N(\mu, \sigma^2)$
- Not $Y = \alpha + \beta X + Z$

Two meanings, two very different use cases.

Inference techniques offering valid p-values and confidence statements under minimal assumptions: sign test, median test, Wilcoxon's tests, and the Kruskal-Wallis and Friedman tests, tests of independence. 2016 Stat 205

Predictive modeling techniques valid quite generally, such as kernel and spline smoothing, nearest neighbor, and even deep neural networks. 2021 Stat 205

What is Nonparametric?, 3

Nonparametric Inference Use Case:

Clinical Medicine An intervention (drug/treatment) will be given to a group of patients.

Inference question Does the intervention change an outcome measure in a statistically significant way?

Prevalence Roughly 1 Million Clinical Medical Articles per year.

Mayo Vaccine Study, 1

Comparison of two highly-effective mRNA vaccines for COVID-19 during periods of Alpha and Delta variant prevalence

Arjun Puranik[1+], Patrick J. Lenehan[1+], Eli Silvert[1], Michiel J.M. Niesen[1], Juan Corchado-Garcia[1], John C. O'Horo[2], Abinash Virk[2], Melanie D. Swift[2], John Halamka[2], Andrew D. Badley[2], A.J. Venkatakrishnan[1], Venky Soundararajan[1]

[1] nference, Cambridge, Massachusetts 02139, USA

[2] Mayo Clinic, Rochester, Minnesota 55902, USA

[+] These authors contributed equally

* Correspondence to: Venky Soundararajan (venky@nference.net)

Abstract

Although clinical trials and real-world studies have affirmed the effectiveness and safety of the FDA-authorized COVID-19 vaccines, reports of breakthrough infections and persistent emergence of new variants highlight the need to vigilantly monitor the effectiveness of these vaccines. Here we compare the effectiveness of two full-length Spike protein-encoding mRNA vaccines from Moderna (mRNA-1273) and Pfizer/BioNTech (BNT162b2) in the Mayo Clinic Health System over time from January to July 2021, during which either the Alpha or Delta variant was highly prevalent. We defined cohorts of vaccinated and unvaccinated individuals from Minnesota (n = 25,589 each) matched on age, sex, race, history of prior SARS-CoV-2 PCR testing, and date of full vaccination. Both vaccines were highly effective during this study period against SARS-CoV-2 infection (mRNA-1273: 86%, 95%CI: 81-90.6%; BNT162b2: 76%, 95%CI: 69-81%) and COVID-19

Mayo Vaccine Study, 2

Methods

Study design and population

This is a retrospective study of individuals who underwent SARS-CoV-2 polymerase chain reaction (PCR) testing at the Mayo Clinic and hospitals affiliated with the Mayo Clinic Health System (Arizona, Florida, Iowa, Minnesota, and Wisconsin). Overall, there were 645,109 individuals with at least one SARS-CoV-2 PCR test. We included individuals who met the following criteria: (i) age greater than or equal to 18 years; (ii) received at least one dose of BNT162b2 or mRNA-1273 after December 1, 2020 and on or before July 29, 2021; (iii) did not have any positive SARS-CoV-2 PCR tests prior to their first vaccine dose; and (iv) did not receive a mismatched series of COVID-19 vaccines (e.g., did not receive doses from more than one manufacturer). There were 119,463 individuals who met these criteria for BNT162b2 and 60,083 individuals who met these criteria for mRNA-1273.

Mayo Vaccine Study, 3

Defining clinical outcomes of interest

To perform overall and comparative analyses of vaccine effectiveness, the following outcomes were assessed for each cohort:

1. *SARS-CoV-2 infection*: at least one positive SARS-CoV-2 PCR test. The date of infection was taken as the date of the first positive test.
2. *COVID-19 associated hospitalization*: admission to the hospital occurring within 21 days after SARS-CoV-2 infection.
3. *COVID-19 associated ICU admission*: admission to the intensive care unit (ICU) occurring within 21 days after SARS-CoV-2 infection.
4. *COVID-19 associated mortality*: death occurring within 28 days after SARS-CoV-2 infection.
5. *Breakthrough infection*: a SARS-CoV-2 infection occurring after full vaccination (i.e., at least 14 days after the second dose of mRNA-1273 or BNT162b2).

Mayo Vaccine Study, 4

Results

From January to July 2021 in Minnesota, the effectiveness estimates of mRNA-1273 and BNT162b2 in preventing SARS-CoV-2 infection with onset at least 14 days after the second dose were 86% (95% CI: 81-90.6%, $p=1.6 \times 10^{-42}$) and 76% (95% CI: 69-81%, $p=1.3 \times 10^{-31}$), respectively (**Figure 1**, **Table 2**, **Figure S2A**). Full vaccination with either vaccine was also highly effective against COVID-19 associated hospitalization (mRNA-1273: 91.6%, 95% CI: 81-97%, $p=8.3 \times 10^{-14}$; BNT162b2: 85%, 95% CI: 73-93%, $p=3.8 \times 10^{-12}$), ICU admission (mRNA-1273: 93.3%, 95% CI: 57-99.8%, $p=5.0 \times 10^{-4}$; BNT162b2: 87%, 95% CI:46-98.6%, $p=1.2 \times 10^{-3}$), and death (no deaths in either cohort) (**Table 2**, **Figure S2B-C**).

Mayo Vaccine Study, 5

These estimates of effectiveness against infection (86% and 76%) were lower than those that we previously observed in the Mayo Clinic Health System through April 20, 2021 (mRNA-1273: 93.3%, 95% CI: 85.7-97.4%; BNT162b2: 86.1%, 95% CI: 82.4-89.1%).[6] We thus analyzed the effectiveness of full vaccination longitudinally on a monthly basis starting in February 2021 (see **Methods**). In the context of increasing cases in Minnesota during July (**Figure S3**), the effectiveness against infection was lower for mRNA-1273 (76%, 95% CI: 58-87%) compared to prior months, with an even more pronounced reduction for BNT162b2 (42%, 95% CI: 13-62%) (**Figure 2A**; **Table 3**). Importantly, the effectiveness of mRNA-1273 and BNT162b2 against COVID-19 associated hospitalization has remained more consistently high (**Figure 2B**, **Table 4**). Of note, July corresponds to the time during which the Delta variant has risen to prominence in Minnesota (**Figure 2C**).

Mayo Vaccine Study, 6

In addition to the changing effectiveness against infection over time, we noted that the 95% confidence intervals of the estimates for effectiveness against infection did not overlap (mRNA-1273: 81-90.6%; BNT162b2: 69-81%) (**Table 2**). The incidence rate of breakthrough infections over the study duration was significantly lower in the mRNA-1273 cohort ($IR_{mRNA\text{-}1273}$: 0.017, $IR_{BNT162b2}$: 0.031; IRR = 0.56, 95% CI: 0.36-0.83) despite similar baseline infection risks in the week after the first vaccine dose (IRR = 1.3; 95% CI: 0.89-1.8) (**Table 2**). Kaplan-Meier analysis indicates a difference in cumulative breakthrough infection incidence between the vaccinated cohorts ($p=3.4x10^{-3}$; **Figure S2A**). On the other hand, the mRNA-1273 and BNT162b2 cohorts had similar rates of hospitalization (IRR: 0.57, 95% CI: 0.17-1.7, p=0.30), ICU admission (IRR: 0.53, 95% CI: 0.0089-10, p=0.59), and death (no events in either cohort) (**Table 2** and **Figures S2B-C**).

Inference in Publications

- Inference: *the* deliverable published research must provide.
- 'User Interface' for readers
  - *P*-values
  - Confidence Intervals
- Often the *only* interface reported widely.

Parametric Inference

- Normal Model
$$X_i \sim_{iid} N(\theta, \sigma^2), \qquad i = 1, \ldots, n.$$
- Null Hypothesis: $H_0 : \theta = 0$; Alternative Hypothesis: $H_1 : \theta \neq 0$
- Test statistic: (t-test)
$$T = \sqrt{n} \cdot \frac{\bar{X}_n}{s_n}$$
- Rejection region $|T| > t_{n-1,1-\alpha/2}$;
- $t_{n-1,1-\alpha/2}$ avail. R; large-n $t_{n-1,1-\alpha/2} \sim \mathfrak{Z}_{1-\alpha/2}$
Convention: ($\mathfrak{Z}_u \equiv \Phi^{-1}(u)$; in R: `qnorm(u)`)
- Assumptions: normally distributed data
- Failure Modes: outliers (especially)
- Benefits: if assumptions are true, *optimal performance*

nonParametric Inference, 1

- nonParametric Model
$$X_i \sim_{iid} F(\cdot - \theta), \qquad i = 1, \ldots, n.$$
F is any zero-median CDF: $P\{X < 0\} = P\{X > 0\}$.
- Null Hypothesis: $H_0 : \theta = 0$; Alternative Hypothesis: $H_1 : \theta \neq 0$
- Test statistic: (*Sign Test*)
$$S = \sum_i \text{sign}(X_i)$$
- Rejection region $|S| > s_{n,\alpha}$
$s_{n,\alpha}$ available using R's `pbinom`
Large-n: $s_{n,\alpha} \approx \mathfrak{Z}_{1-\alpha/2} \cdot \sqrt{n}$
- Assumptions: F has zero median.
- Failure Modes: asymmetry, dependence
- Benefits:
  - Distribution-free type I error probability
  - total immunity to outliers
- Questions: what about *performance*?

nonParametric Inference, 2

- nonParametric Model
$$X_i \sim F(\cdot - \theta), \qquad i = 1, \ldots, n.$$
F is any symmetric CDF: $P\{X < -x\} = P\{X > x\}$, $\forall x$.
- Null Hypothesis: $H_0 : \theta = 0$; Alternative Hypothesis: $H_1 : \theta \neq 0$
- Test statistic: (*Wilcoxon Signed-Rank*)
$$W = \sum_i \text{sign}(X_i) R|X_i|$$
$R|X_i|$ = rank of $|X_i|$ among $\{|X_1|, \ldots, |X_n|\}$.
- Rejection region $|W| > w_{n,\alpha}$
- $w_{n,\alpha}$ available from R;
Large-n: $w_{n,\alpha} \approx \mathfrak{Z}_{1-\alpha/2} \cdot \sqrt{n(n+1)(2n+1)/6}$;
- Assumptions: symmetry
- Failure Modes: asymmetry, dependence
- Benefits:
  - distribution-free type I error probability
  - total immunity to outliers
- Questions: what about *performance*?

nonParametric Inference, 3

- nonParametric Model
$$X_i \sim F(\cdot - \theta), \qquad i = 1, \ldots, n.$$
F is any symmetric CDF: $P\{X < -x\} = P\{X > x\}$, $\forall x$.
- Null Hypothesis: $H_0 : \theta = 0$; Alternative Hypothesis: $H_1 : \theta \neq 0$
- Test statistic: (*Normal Scores*)
$$T_{ns} = \sum_i \text{sign}(X_i) a_{ns}(R|X_i|)$$
- $a_{ns}(i; n) = \mathfrak{Z}_{(1+\frac{i}{n+1})/2}$; `qnorm( (1 + i/(n+1))/2 )` (in R).
- Small-n rejection region $|T_{ns}| > t^{ns}_{n,\alpha}$ available by simulation;
Large-n: $t^{ns}_{n,\alpha} \approx \mathfrak{Z}_{1-\alpha/2} \cdot \sqrt{n}$
- Assumptions: symmetry
- Benefits:
  - distribution-free type I error probability
  - total immunity to outliers
- Questions: what about *performance*?

Comparing all Four Tests, 1

Name	Statistic
t-test	$T = \sqrt{n} \cdot \frac{\bar{X}_n}{s_n}$
Sign-test	$S = \sum_i \text{sign}(X_i)$
Wilcoxon Signed-Rank	$W = \sum_i \text{sign}(X_i) R\|X_i\|$
Normal Scores	$T_{ns} = \sum_i \text{sign}(X_i) a_{ns}(R\|X_i\|)$

$$R|X_i| = \text{rank of } |X_i| \text{ among } \{|X_1|, \ldots, |X_n|\}.$$

$$R|X_i| \equiv \sum_j 1_{\{X_j \le X_i\}}.$$

Comparing all Four Tests, 2

Name	Statistic	Null Hypothesis
t-test	$T = \sqrt{n} \cdot \frac{\bar{X}_n}{s_n}$	$\{N(0, \sigma^2) : \sigma > 0\}$
Sign-test	$S = \sum_i \text{sign}(X_i)$	$\{F : F(0) = 1/2\}$
Wilcoxon	$W = \sum_i \text{sign}(X_i) R\|X_i\|$	$\{F : \text{symmetric about } 0\}$
Normal Scores	$T_{ns} = \sum_i \text{sign}(X_i) a_{ns}(R\|X_i\|)$	$\{F : \text{symmetric about } 0\}$

Null Hypothesis:
The collection of all distributions obeying the hypothesis H_0.
Can be small or large.

$\mathcal{F}_{norm}$	mean zero normal:	$N(0, \sigma^2) : \sigma > 0$
$\mathcal{F}_{sym}$	symmetric about 0:	$F(a, b) = F(-b, -a)$ for all $a < b$
$\mathcal{F}_{zmed}$	median 0	$F(0) = 1/2$.

$$\mathcal{F}_{norm} \subset \mathcal{F}_{sym} \subset \mathcal{F}_{zmed}$$

Comparing all Four Tests, 3

Name	Statistic	Outlier Immunity
t-test	$T = \sqrt{n} \cdot \frac{\bar{X}_n}{s_n}$	No
Sign-test	$S = \sum_i \text{sign}(X_i)$	Yes
Wilcoxon	$W = \sum_i \text{sign}(X_i) R\lvert X_i \rvert$	Yes
Normal Scores	$T_{ns} = \sum_i \text{sign}(X_i) a_{ns}(R\lvert X_i \rvert)$	Yes

Outlier immunity:
no matter if some $|X_i|$ are anomalously large –
the raw values of $|X_i|$ don't appear in any sums!

Comparing all Four Tests, 4

Name	Statistic	Distribution-Free
t-test	$T = \sqrt{n} \cdot \frac{\bar{X}_n}{s_n}$	No
Sign-test	$S = \sum_i \text{sign}(X_i)$	Yes
Wilcoxon	$W = \sum_i \text{sign}(X_i) R\lvert X_i \rvert$	Yes
Normal Scores	$T_{ns} = \sum_i \text{sign}(X_i) a_{ns}(R\lvert X_i \rvert)$	Yes

Distribution-Free:
The test critical value is the same across distinct members of H_0
The type-1 error of the test is the same across distinct members of H_0

Comparing all Four Tests, 5

Name	Statistic	Performance at $N(0,1)$
t-test	$T = \sqrt{n} \cdot \frac{\bar{X}_n}{s_n}$	Optimal
Sign-test	$S = \sum_i \text{sign}(X_i)$	Suboptimal
Wilcoxon	$W = \sum_i \text{sign}(X_i) R\lvert X_i\rvert$	Suboptimal
Normal Scores	$T_{ns} = \sum_i \text{sign}(X_i) a_{ns}(R\lvert X_i\rvert)$	(Asy.) Optimal

Optimality:
Maximal Power at a given fixed level α.
Smallest number of samples needed to attain a given power.

Generalizing The Rank Tests, 1

Name	Statistic
Sign-test	$S = \sum_i \text{sign}(X_i)$
Wilcoxon	$W = \sum_i \text{sign}(X_i) R\lvert X_i\rvert$
Normal Scores	$T_{ns} = \sum_i sgn(X_i) a_{ns}(R\lvert X_i\rvert)$
Rank-Score	$T_\phi = \sum_i sgn(X_i) a_\phi(R\lvert X_i\rvert)$

$$a_\phi(r) = \phi(\frac{r}{n+1}).$$

where $\phi(u)$ some score generating function

Generalizing The Rank Tests, 2

Name	Statistic	Score generator
Sign-test	$S = \sum_i \text{sign}(X_i)$	$\phi(u) = 1$
Wilcoxon	$W = \sum_i \text{sign}(X_i) R\lvert X_i\rvert$	$\phi(u) = u$
Normal Scores	$T_{ns} = \sum_i \text{sign}(X_i) a_{ns}(R\lvert X_i\rvert)$	$\phi(u) = 3_{(1+u)/2}$
Rank-Score	$T_\phi = \sum_i \text{sign}(X_i)\phi(R\lvert X_i\rvert)$	

Parametric Two-Sample Inference

- Normal Shift Model, equal variance

$$X_i \sim_{iid} N(\theta_X, \sigma^2), \qquad i = 1, \ldots, n_X.$$

$$Y_i \sim_{iid} N(\theta_Y, \sigma^2), \qquad i = 1, \ldots, n_Y.$$

- Null Hypothesis: $H_0 : \theta_X = \theta_Y$; Alternative Hypothesis: $H_1 : \theta_X = \theta_Y$
- Test statistic: (t-test, pooled variance)

$$T = \sqrt{n^+} \cdot \frac{\bar{X}_{n_1} - \bar{Y}_{n_2}}{s^+}$$

- $n^+ = (n_X - 1) + (n_Y - 1)$; $(s^+)^2 = ((n_X - 1)s_X^2 + (n_Y - 1)s_Y^2)/n^+$.
- Rejection region $|T| > t_{n^+, 1-\alpha/2}$;
- $t_{n^+, 1-\alpha/2}$ avail. R;
  Large-n^+ $t_{n^+, 1-\alpha/2} \sim \mathfrak{Z}_{1-\alpha/2}$

- Assumptions: normally distributed data
- Failure Modes: outliers (especially)
- Benefits: if assumptions are true, *optimal performance*

nonParametric Two-Sample Inference, 1

- nonParametric Model

$$X_i \sim_{iid} F, \qquad i = 1, \ldots, n.$$

$$Y_i \sim_{iid} G, \qquad i = 1, \ldots, n.$$

  F, G any CDFs.
- Matched Pairs: $n = n_X = n_Y$, items 'matched' in some way.
- Null Hypothesis: $H_0 : F = G$; Alternative Hypothesis: $H_1 : F \lesssim G$
- Test statistic: (*Sign Test*)

$$S = \sum_i \text{sign}(X_i - Y_i)$$

- Rejection region $|S| > s_{n,\alpha}$
- Assumptions: (X_i, Y_i) independent and identically distributed.
- Failure Modes: asymmetry, dependence
- Benefits:
  - Distribution-free type I error probability
  - total immunity to outliers
- Questions: what about *performance*?

nonParametric Two-Sample Inference, 2

- nonParametric Model

$$X_i \sim_{iid} F, \qquad i = 1, \ldots, n_X.$$

$$Y_i \sim_{iid} G, \qquad i = 1, \ldots, n_Y.$$

  F, G any CDFs.
- Null Hypothesis: $H_0 : F = G$; Alternative Hypothesis: $H_1 : F \lesssim G$
- Test statistic: (*Wilcoxon Rank-Sum*)

$$W = \sum_i R(Y_i)$$

  $R(Y_i) = R(Y_i; X \cup Y)$ rank of Y_i among $\{X_1, \ldots, X_{n_X}\} \cup \{Y_1, \ldots, Y_{n_Y}\}$.
  Aka *Mann-Whitney*

$$U = \sum_{i,j} 1_{\{Y_i \geq X_j\}} = (W - \frac{n_Y(n_Y+1)}{2})$$

- Rejection region $|W - E_0 W| > w_{n,\alpha}$
  $w_{n,\alpha}$ available from *R* `wilcox.test`
  $E_0 W = (n_X + n_Y)/2$ (expected rank of Y_1 under H_0)
- Assumptions: Null=No change in distribution.
- Benefits:
  - distribution-free type I error probability
- Questions: what about *performance*?

nonParametric Two-Sample Inference, 3

- nonParametric Model

$$X_i \sim_{iid} F, \qquad i = 1, \ldots, n_X.$$

$$Y_i \sim_{iid} G, \qquad i = 1, \ldots, n_Y.$$

  F, G any CDFs.
- Null Hypothesis: $H_0 : F = G$; Alternative Hypothesis: $H_1 : F \lesssim G$
- Test statistic: (*Normal Scores*)

$$T_{ns} = \sum_i a_{ns}[R(Y_i)]$$

- $a_{ns}(i; n) = \mathfrak{Z}_{\frac{i}{n+1}}$; `qnorm( i/(n+1) )` (in R).
- Rejection region $|T_{ns}| > t^{ns,2}_{n,\alpha}$
  Small-n: $t^{ns,2}_{n,\alpha}$ available by simulation
  Large-n: $t^{ns,2}_{n,\alpha} \approx \mathfrak{Z}_{1-\alpha/2} \cdot \sqrt{n}$
- Assumptions: Null=no change in distribution
- Benefits:
  - distribution-free type I error probability
- Questions: what about *performance*?

Comparing all Four Two-Sample Tests, 1

Name	Statistic
t-test	$T = \sqrt{n^+} \cdot \frac{\bar{X}_{n_X} - \bar{Y}_{n_Y}}{s^+}$
Wilcoxon Rank-Sum	$W = \sum_i R(Y_i)$
Normal Scores	$T_{ns} = \sum_i a_{ns}[R(Y_i)]$
Sign-test Matched Pairs	$S = \sum_i \text{sign}(X_i - Y_i)$

$R|X_i| = \text{rank of } |X_i| \text{ among } \{|X_1|, \ldots, |X_n|\}.$

$$R|X_i| \equiv \sum_j 1_{\{X_j \leq X_i\}}.$$

Comparing all Four Two-Sample Tests, 2

Name	Statistic	Null Hypothesis
t-test	$T = \sqrt{n^+} \cdot \frac{\bar{X}_{n_X} - \bar{Y}_{n_Y}}{s^+}$	$\{(\mu_X = \mu_Y, \sigma_X = \sigma_Y)\}$
Wilcoxon Rank-Sum	$W = \sum_i R(Y_i)$	$F = G$
Normal Scores	$T_{ns} = \sum_i a_{ns}[R(Y_i)]$	$F = G$
Sign-test Matched Pairs	$S = \sum_i \text{sign}(X_i - Y_i)$	$F = G$

Null Hypothesis:
The collection of all pairs distributions obeying the hypothesis H_0.
Collection can be small or large.

$\mathcal{F}_{norm}$ equi-Normal: $F_X = F_Y = N(\mu, \sigma^2) : \mu \in \mathbf{R}, \sigma > 0$
$\mathcal{F}_{arb}$ equi-Arb: $F_X = F_Y = F$ arbitrary

$$\mathcal{F}_{norm} \subset \mathcal{F}_{arb}$$

Comparing all Four Two-Sample Tests, 3

Name	Statistic	Outlier Immunity
t-test	$T = \sqrt{n^+} \cdot \frac{\bar{X}_{n_X} - \bar{Y}_{n_Y}}{s^+}$	No
Wilcoxon Rank-Sum	$W = \sum_i R(Y_i)$	Yes
Normal Scores	$T_{ns} = \sum_i a_{ns}[R(Y_i)]$	Yes
Sign-test Matched Pairs	$S = \sum_i \text{sign}(X_i - Y_i)$	Yes

Outlier immunity:
no matter if some X_i or Y_i are anomalously large –
the raw values of X_i and Y_i don't appear in any formulas!

Comparing all Four Two-Sample Tests, 4

Name	Statistic	Distribution-Free
t-test	$T = \sqrt{n^+} \cdot \frac{\bar{X}_{n_X} - \bar{Y}_{n_Y}}{s^+}$	No
Wilcoxon Rank-Sum	$W = \sum_i R(Y_i)$	Yes
Normal Scores	$T_{ns} = \sum_i a_{ns}[R(Y_i)]$	Yes
Sign-test Matched Pairs	$S = \sum_i \text{sign}(X_i - Y_i)$	Yes

Distribution-Free:
The test critical value is the same across distinct members of H_0
The type-1 error of the test is the same across distinct members of H_0

Comparing all Four Two-Sample Tests, 5

Name	Statistic	Performance at $N(0,1)$
t-test	$T = \sqrt{n^+} \cdot \frac{\bar{X}_{n_X} - \bar{Y}_{n_Y}}{s^+}$	Optimal
Wilcoxon Rank-Sum	$W = \sum_i R(Y_i)$	Suboptimal
Normal Scores	$T_{ns} = \sum_i a_{ns}[R(Y_i)]$	(Asy.) Optimal
Sign-test Matched Pairs	$S = \sum_i \text{sign}(X_i - Y_i)$	Suboptimal

Optimality:
Maximal Power at a given fixed level α.
Smallest number of samples needed to attain a given power.

Quantifying Performance

Name	Statistic	$e_{\cdot \mid t}(N(0,1))$	$\min_F e_{\cdot \mid T}(F)$
t-test	$T = \sqrt{n^+} \cdot \frac{\bar{X}_{n_X} - \bar{Y}_{n_Y}}{s^+}$	1	1
Wilcoxon Rank-Sum	$W = \sum_i R(Y_i)$	0.955	0.864
Normal Scores	$T_{ns} = \sum_i a_{ns}[R(Y_i)]$	1	1
Sign-test Matched Pairs	$S = \sum_i \text{sign}(X_i - Y_i)$	$2/\pi$	$2/\pi$

Asymptotic Relative Efficiency:
Ratio of sample sizes needed to achieve given power:

$$e_{T_1 \mid T_0} = \text{“}lim\text{”} \quad \frac{n_{T_0}}{n_{T_1}}$$

(Quite involved mathematically to explain the definition precisely, but heuristically clear kind of miraculous that simple clear answers exist)

Example 3.2.1 Esophageal Cancer, 1

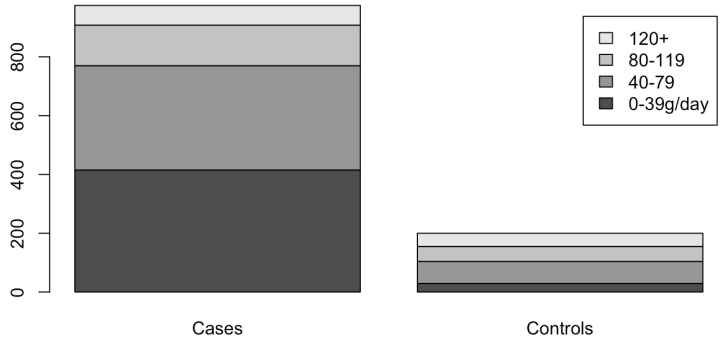
(Breslow et al. 1980)

- Case-control study of esophageal cancer.
- Null Hypothesis: Alcohol consumption same in the two groups.

```
library(datasets); data(esoph); head(esoph)
```

```
##   agegp     alcgp    tobgp ncases ncontrols
## 1 25-34 0-39g/day 0-9g/day      0        40
## 2 25-34 0-39g/day    10-19      0        10
## 3 25-34 0-39g/day    20-29      0         6
## 4 25-34 0-39g/day      30+      0         5
## 5 25-34     40-79 0-9g/day      0        27
## 6 25-34     40-79    10-19      0         7
```

Example 3.2.1 Esophageal Cancer, 2




```
wilcox.test(x,y,alternative = "greater")
```

```
##
##  Wilcoxon rank sum test with continuity correction
##
## data:  x and y
## W = 135610, p-value < 2.2e-16
## alternative hypothesis: true location shift is greater than 0
```

Reject H_0 alcohol consumption is the same