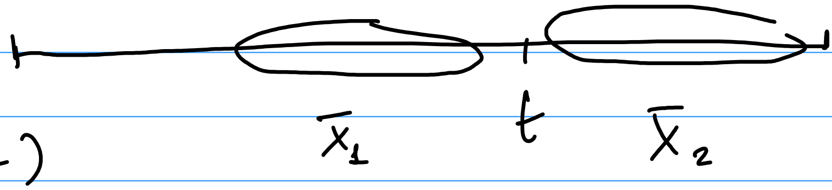
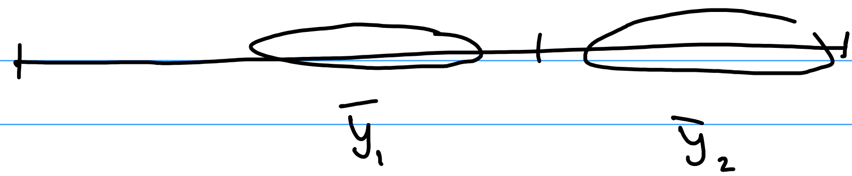
A / B Testing

1) Time Effects

A (treatment)

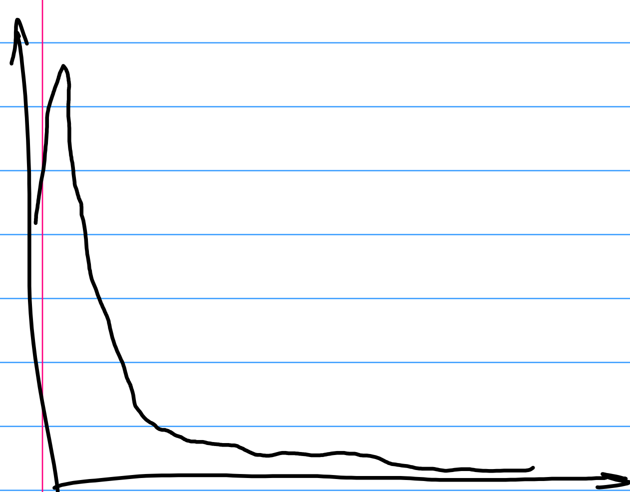


B (control)



$\bar{x}_1 = \bar{y}_1$ $\quad$ $\bar{x}_2 \overset{?}{=} \bar{y}_2$

2) Skewed



⊖ Remove outliers

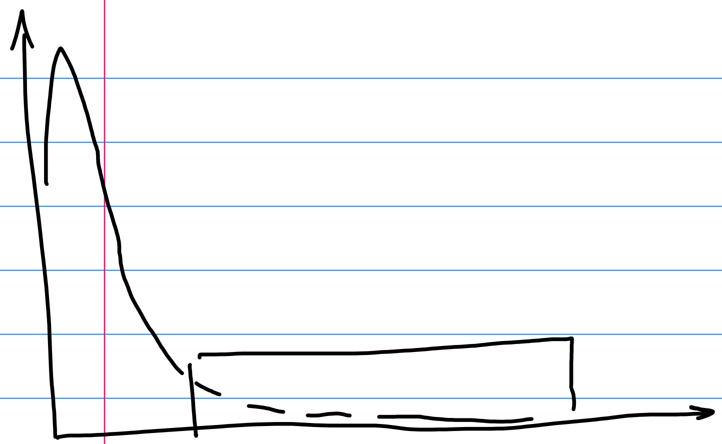
⊕ Transformation

- logarithmic

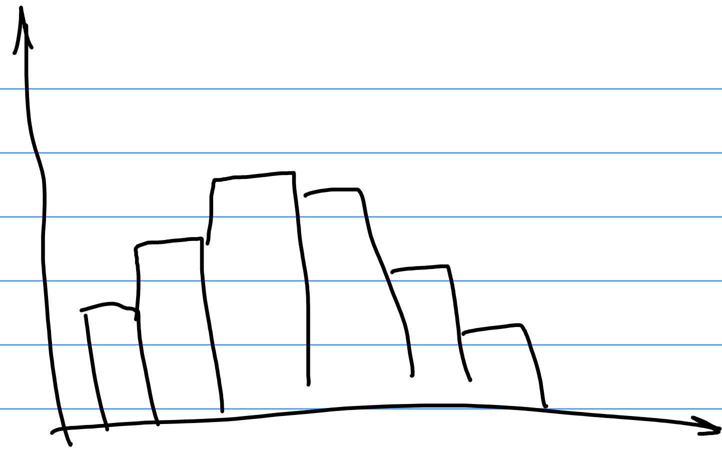
(Test for normality)

- Box-Cox

3) Too Big Data



Bucketing



$$x_j^B = \sum_i^{N_j} x_{ji}$$

N_j – # obs in bucket j

$$\frac{s^2}{N} \approx \frac{s_B^2}{B}$$

B — # buckets

Metrics

1) User-level $U = \dfrac{\sum X_i}{N}$ $\qquad \dfrac{3 + \dots + 4}{1 + \dots + 1}$

2) Ratio $R = \dfrac{\sum X_i}{\sum Y_i}$ $\qquad \dfrac{2 + \dots + 5}{10 + \dots + 6}$

X, Y, R - r.v.

1) Bootstrap

2) Linearization

Taylor expansion of X/Y

at $(E(X), E(Y))$

$$Z = \frac{E[X]}{E[Y]} + \frac{1}{E[Y]} \cdot \left(X - \frac{E[X]}{E[Y]} \cdot Y\right)$$

$$E(Z) \approx E(R) \qquad Var(Z) \approx Var(R)$$

3) Delta-method

Hypothesis

		True	False
Decision	Don't reject	TP (confidence) $1-\alpha$	FN (Type II error) β
	Reject	FP (Type I error) α	TN (Power) $1-\beta$

t-test

- normality
- $\delta_x = \delta_y$

$n_x = n_y = n$

$$t = \frac{\bar{X} - \bar{y}}{S_p \cdot \sqrt{2/n}} \sim t_{2n-2}$$

$$S_p = \sqrt{\frac{S_x^2 + S_y^2}{2}}$$

$n_x \neq n_y$

$$t = \frac{\bar{X} - \bar{y}}{S_p^2 \sqrt{\frac{1}{n_x} + \frac{1}{n_x}}} \sim t_{n_x + n_y - 2}$$

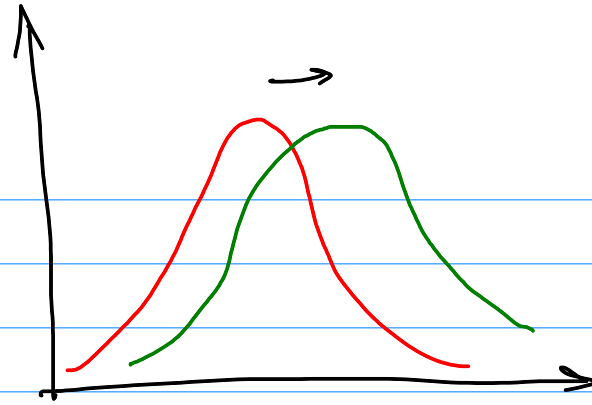
Welch Test

- normality
- $\delta_x \neq \delta_y$

$$t = \frac{\bar{x} - \bar{y}}{\sqrt{\frac{S_x}{n_x} + \frac{S_y}{n_y}}} \sim t_{\hat{d}}$$

Mann-Whitney
(non-parametric)

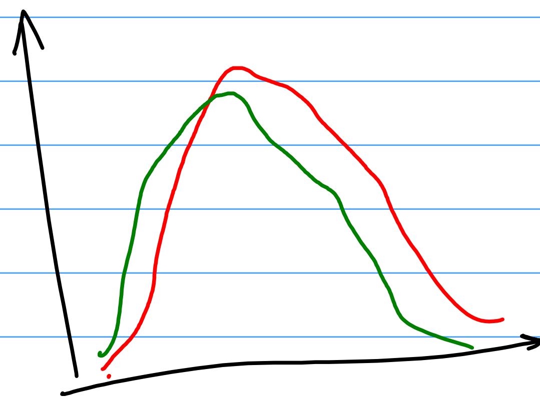
- no assumptions about distr. of X, Y



Test: Is there a difference in rank sum

$H_0: f_X(x) = f_Y(y)$

$H_a: f_X(x) = f_Y(y+a)$



Exact:

1) $U_X = R_X - \frac{n_X(n_X+1)}{2}$

R_i - sum of ranks for

$U_Y = R_Y - \frac{n_Y(n_Y+1)}{2}$

2) $\min[U_X; U_Y]$

Asymptotic:

$U = \sum\sum S(X_i, Y_j)$

$S(X_i, Y_j) = \begin{cases} 1, & X_i > Y_j \\ 0, & X_i < Y_j \end{cases}$

$$z = \frac{U - m_u}{\sigma_u};$$

$$m_u = \frac{n_1 n_2}{2}.$$

$$\sigma_u = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$$

Levene's Test

$H_0: \sigma_1 = \sigma_2 = \dots = \sigma_k$

$H_a: \exists\, i,j \quad \sigma_i \neq \sigma_j$

$$W = \frac{\text{between variance}}{\text{within variance}} = \frac{(N-k) \cdot \sum N_i (\bar{Z}_{i\cdot} - \bar{Z}_{\cdot\cdot})^2}{(k-1) \cdot \sum\sum (Z_{ij} - \bar{Z}_{i\cdot})^2} \sim F(k-1, N-k)$$