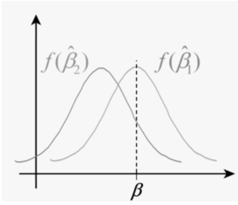
Point Estimators

i. Linear estimator: if is linear function of sample obs

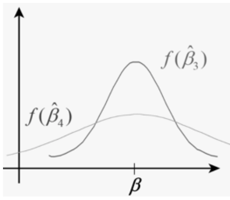
- Ex: $\bar{X}$ is linear estimator of μ : $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$
- Ex: s^2 is quadratic estimator of σ^2 ⟶ doesn't have linear estimator

ii. Unbiased estimator: if $E(\hat{\beta}_1)=\beta$



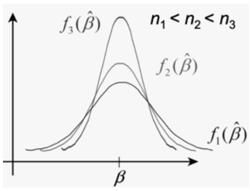
- Biased: $E(\hat{\beta}_2)b \neq \beta$
- Ex: $\bar{X}$ is unbiased estimator of μ: $E(\bar{X}) = \mu$
- Ex: s^2 is unbiased estimator of σ^2

iii. Efficient estimator (within a class e.g. linear): if variance ⩽ variance of other estimator



- Both $\hat{\beta}_3$ and $\hat{\beta}_4$ unbiased but $\hat{\beta}_3$ efficient bc $\hat{\beta}_3$ var s.d. ⩽ $\hat{\beta}_4$ var s.d.
- Ex. $\bar{X}$ (samp mean) is efficient (best) estimator of μ *in* all linear unbiased estimator of μ

iv. Consistent estimator: if samp distribution turn into straight line at pop parameter β when $n \to \infty$, var ⟶ 0, $E(\beta) \to \beta$



Parametric vs Non-parametric

- Parametric: test pop μ
  - Assume variable quantitative on interval/ratio scale
  - Assume: population normality
    - pop s.d. unknown, pop normal distributed => t
    - pop s.d. known, samp mean normal => Z
  - Assume random sampling
- Non-parametric: test pop η
  - No assume pop normality
  - With weaker assumptions (when assumption not satisfied)

Non-Parametric Test Pop Central Location η

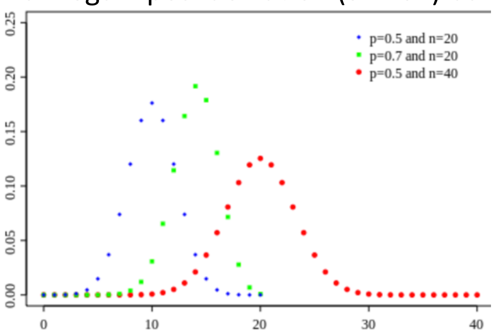
- Mean: comprehensive consider all value, mean only exist in interval/ratio
- Median: not affected by outlier, can do ordinal bc median exist

#1 ***Sign Test***

- Special case of binomial test p=0.5
- Assumptions:
  - data is random sample
  - variable interested is continuous and at least ordinal scale (or discrete with large possibility)
- H0: $\eta = \eta_0$, HA: $\eta \neq > < \eta_0$
- Procedure:
  1) Test stat: S (S- or S+ use S+)
     - S+: # positive deviations from η_0
       - No include $X_i = \eta_0$
     - S- + S+ = n*

DT	Median under H0	Deviation
7	6	1
3	6	-3
4	6	-2
6	6	0
10	6	4
5	6	-1
6	6	0
4	6	-2
3	6	-3
8	6	2

  2) P-value: no CV
     - n*⩽25: Binomial distribution with n (# non-zero deviation) and p=0.5
       - Table 1
       - $S \sim B(n, 0.5)$ ⟶ $E(S) = 0.5n$, $Var(S) = 0.25n$ Var = npq
       - H0: nega = posi deviation (S- = S+) bc med in center



       - $p_R = P(S \geqslant S+)$, or $p_L = P(S \leqslant S+)$, or $2*\min(p_R, p_L)$
     - n*> 25: approx. Normal distribution if np=nq=0.5⩾5, n⩾10
       - Table 3
       - $B(n, p) \sim N(np, \sqrt{npq})$ ⟶ $S \sim N(0.5n, 0.5\sqrt{n})$ ***
       - N(mean, sd)
       - p_R : $P(S \geq S^+) = P(S^+ \leq S \leq n*)$

$$\approx P\left(Z < \frac{(n*-0.5)-0.5n*}{0.5\sqrt{n*}}\right) - P\left(Z < \frac{(S^+ + 0.5)-0.5n*}{0.5\sqrt{n*}}\right)$$

- p_L : $P(S \le S^+) = P(0 \le S \le S^+)$

$$\approx P\left(Z < \frac{(S^+ + 0.5)-0.5n*}{0.5\sqrt{n*}}\right) - P\left(Z < \frac{(0+0.5)-0.5n*}{0.5\sqrt{n*}}\right)$$

  - ± 0.5 to make probability wider

3) Reject H0 if
   i) right-tail: $p_R = P(S \geqslant S+) < \alpha$
   ii) left-tail: $p_L = P(S \leqslant S+) < \alpha$
   iii) two-tail: $2*\min(p_R, p_L) < \alpha$ ⟶ whichever p_R p_L is smaller

- R: binom.test (when #success known), SignTest (when #success unknown) from DescTools
  - Exact binomial test, One-sample Sign-Test

#2 ***Wilcoxon Signed Ranks Test***

- Assumptions:
  - data is random sample
  - variable interested is continuous and interval/ratio scale
  - distribution of pop is symmetric (μ=η)
- H0: $\eta = \eta_0$, HA: $\eta \neq > < \eta_0$
- Procedure:
  1) **Rank** all absolute non-zero **deviations** $|d_i|$ small to large, note their signs

DT	Median under H0	Deviation	Abs. deviation	Rank
7	6	1	1	1.5
3	6	-3	3	6.5
4	6	-2	2	4.0
6	6	0	0	
10	6	4	4	8.0
5	6	-1	1	1.5
6	6	0	0	
4	6	-2	2	4.0
3	6	-3	3	6.5
8	6	2	2	4.0

tie: $\sum \text{rank} \div \text{amount}$

  2) **Sum** rank of T- and T+ , $T^- + T^+ = \frac{n(1+n)}{2}$
  3) Test stat: $T^- = T^+ = \frac{n(1+n)}{4}$
     - $n* \leqslant 30$: Table 9
       - Give T_L T_U
       - Test stat: $T_{obs} = T+$
       - Reject H0 if
         i) right-tail: $T \geqslant T_{U,\alpha}$
         ii) left-tail: $T \leqslant T_{L,\alpha}$
         iii) two-tail: $T \geqslant T_{U,\alpha/2}$ or $T \leqslant T_{L,\alpha/2}$
     - $n* > 30$: Table 3 approx Normal
       - $T \sim N(\mu_T, \sigma_T)$, $\mu_T = \frac{n(n+1)}{4}$, $\sigma_T^2 = \frac{n(n+1)(2n+1)}{24}$
       - Test stat: $Z_{obs} = \frac{T-\mu_T}{\sqrt{\sigma_T{}^2}}$