Grundläggande matematisk statistik

Icke-parametriska test

Uwe Menzel, 2018
uwe.menzel@matstat.org
www.matstat.org

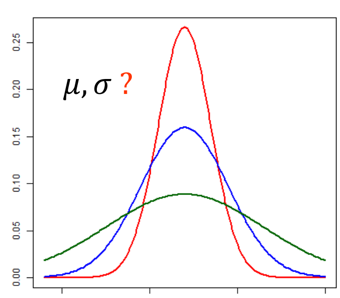
Testmetoder

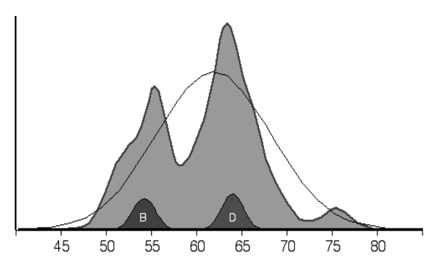
Parametriska metoder

Fördelningen för populationen som stickprovet togs ifrån är känd så nära som på ett antal parametrar, t.ex: N med okända μ och/eller σ

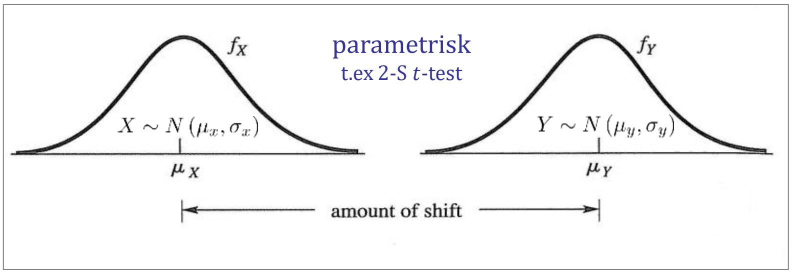
Icke-parametriska metoder

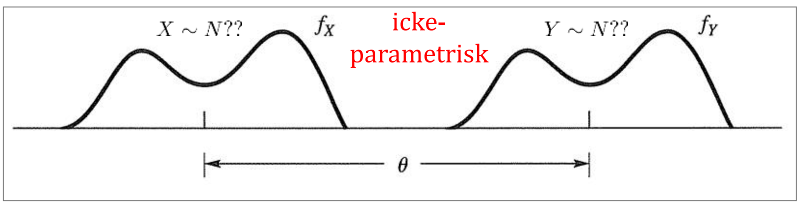
Vill helst ha test med samma förmåga men utan antaganden om den underliggande fördelningen!





”Location shift”





www.matstat.org

När behövs ett icke-parametrisk test ?

- När förutsättningarna för ett parametriskt test inte är uppfyllda:
  - 2-Sample t-test: de underliggande populationerna måste vara normalfördelade
  - One-Way ANOVA: kräver (ungefär) normalfördelning för alla grupper, dessutom (ungefär) homogena varianser
- Om det är svårt eller omöjligt att kvantifiera data:
  - ordinalskalor (storleksordning finns, men differenser saknar betydelse: t.ex. kläderstorlek, betyg i skolan: A, B, C, …)
  - sådana data kan dock rangordnas (ordnas enligt storlek)

Hur pass bra fungerar icke-parametriska test ?

- Nästan så bra som parametriska test när förutsättningarna (normalfördelning, varianshomogenitet) föreligger
  - styrkan (power, $1-\beta$) ofta mindre jämfört med parametriska test
- Är ofta bättre när förutsättningarna (normalfördelning osv.) inte är uppfyllda

www.matstat.org

Vilket test ”ersätts” med vilket ?

population normalfördelad	population inte normalfördelad	
1-Sample t-test (1-Sample Z-test)	1-Sample Sign test	allmän fördelning R: sign.test
	1-Sample Wilcoxon test	symmetrisk fördelning
Paired t-test	Matched Pairs Sign test	1-Sample Sign test på differenserna
	Wilcoxon-Signed Rank test	1-Sample Wilcoxon på differenserna
2-Sample t-test	Mann-Whitney U-test	R: wilcox.test
One-Way ANOVA	Kruskal-Wallis test	R: kruskal.test

www.matstat.org

One-Sample Sign test

- testar om medianen för en population är lika med ett hypotetiskt värde m_0
- Om nollhypotesen stämmer borde ungefär hälften av alla värden ur ett stick-prov vara större än det hypotetiska värdet m_0, andra hälften borde vara mindre.
- Slumpvariabel X: antalet stickprovsvärden som är större än m_0
- n = stickprovsstorlek
- Under H_0 borde X vara binomialfördelad med $p = {}^1/_2$: $X \sim Bin(n, 0.5)$, eftersom för var och en av alla n Bernoulli-försök borde sannolikheten vara 0.5 att stick-provsvärdet är större m_0.

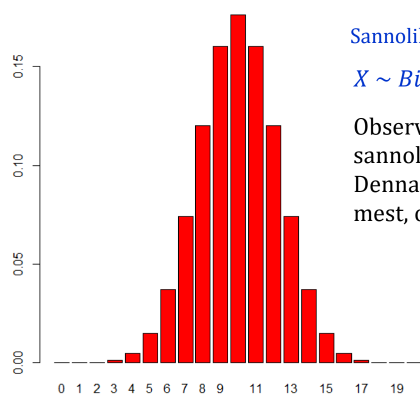
Nollhypoteser och alternativer hypoteser: (x = observation för X)

	Fall 1	**Fall 2**	**Fall 3**
H_0	$m = m_0$	$m < m_0$	$m > m_0$
H_a	$m \neq m_0$	$m > m_0$	$m < m_0$
p-värde	$p = 2 \cdot P(X \leq min(x, n - x))$	$p = P(X \geq x)$	$p = P(X \leq x)$

www.matstat.org

One-Sample Sign Test

- Slumpvariabel X: antalet stickprovsvärden som är större än m_0
- Under H_0 borde X vara binomialfördelad med $p = {}^1/_2$: $X \sim Bin(n, 0.5)$, eftersom för varje Bernoulli-försök borde sannolikheten att stickprovsvärdet är större än m_0 vara 0.5.



Sannolikhetsfunktion

$X \sim Bin(20, 0.5)$

Observationen $x = 10$ är mest sannolik för $n = 2$.
Denna observation förväntar vi oss mest, om nollhypotesen stämmer.

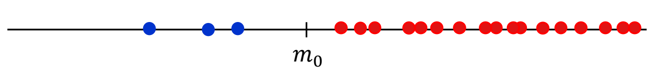
www.matstat.org

One-Sample Sign test

***p*-värde, fall 2** (se tabell ovan): H_0 förkastas om x är stor.

$H_a : m > m_0 \quad \Longrightarrow \quad p = P(X \geq x)$ $\quad \boxed{X = \text{antalet värden} > m_0}$

Exempel: Låt oss anta att vi fick 20 mätvärden ($n = 20$), därav 17 större än den hypotetiska medianen, dvs. $x = 17$.



Det verkar vara osannolikt att den hypotetiska medianen m_0 stämmer.

$p = P(X \geq x) = P(X \geq 17)$ där $X \sim Bin(20, 0.5)$ (”under H_0”)

$p = P(X \geq 17) = 1 - P(X \leq 16) = 1 - F_X(16) = 1 - 0.9987 = \underline{\underline{0.0013}}$

Facit: p-värdet är mindre än $\alpha = 0.05$, vi förkastar H_0, dvs. vi förkastar antagandet att $p = 0.5$ och därmed antagandet att m_0 är medianen för den underliggande fördelningen. Vi accepterar däremot den alternativa hypotesen $m > m_0$ (det sanna värdet på medianen måste vara större än vad hypotesen påstår).

www.matstat.org

Binomialfördelningen, forts.

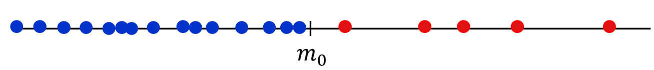
n	$k\backslash p$	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0005	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0036	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0160	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.0510	0.0059
	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.1256	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.2500	0.0577
	7	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.4159	0.1316
	8	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.5956	0.2517
	9	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.7553	0.4119
	10	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.8725	0.5881
	11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9435	0.7483
	12	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9790	0.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

One-Sample Sign test

***p*-värde, fall 3** (se tabell ovan) : H_0 förkastas om x är litet.

$H_a : m < m_0 \Longrightarrow p = P(X \leq x)$ $\boxed{X = \text{antalet värden} > m_0}$

Exempel: Låt oss anta att vi fick 20 mätvärden ($n = 20$), därav 5 större än den hypotetiska medianen, dvs. $x = 5$.



Här verkar det vara osannolikt att den hypotetiska medianen m_0 stämmer.

$p = P(X \leq x) = P(X \leq 5)$ där $X \sim Bin(20, 0.5)$ (”under H_0”)

$p = P(X \leq 5) = F_X(5) = \underline{\underline{0.0207}}$

Facit: p-värdet är mindre än $\alpha = 0.05$, vi förkastar H_0, dvs. vi förkastar antagandet att $p = 0.5$ och därmed antagandet att m_0 är medianen för den underliggande fördelningen. Vi accepterar däremot den alternativa hypotesen $m < m_0$ (det sanna värdet på medianen måste vara mindre än vad hypotesen påstår).

Binomialfördelningen, forts.

n	$k\backslash p$	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0005	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0036	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0160	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.0510	0.0059
	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.1256	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.2500	0.0577
	7	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.4159	0.1316
	8	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.5956	0.2517
	9	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.7553	0.4119
	10	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.8725	0.5881
	11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9435	0.7483
	12	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9790	0.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

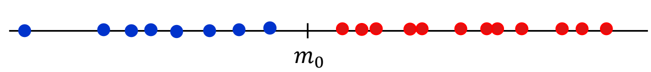
One-Sample Sign test

***p*-värde, fall 1**: H_0 förkastas om x är för litet eller för stort (tvåsidigt test).

$$H_a : m \neq m_0 \quad \Rightarrow \quad p = P\big(X \geq max(x, n-x)\big) + P\big(X \leq min(x, n-x)\big)$$

$$= 2 \cdot P\big(X \leq min(x, n-x)\big)$$ (eftersom $p = 0.5 \rightarrow$ symmetr. fördelning)

<u>Exempel 1</u>: Låt oss anta att vi fick 20 mätvärden ($n = 20$), därav 12 större än den hypotetiska medianen, dvs. $x = 12$.



$$p = P\big(X \geq max(x, n-x)\big) + P\big(X \leq min(x, n-x)\big) \qquad \text{där } X \sim Bin(20, 0.5)$$

$$p = P(X \geq 12) + P(X \leq 8) = 1 - P(X \leq 11) + P(X \leq 8)$$

$$= 1 - F_X(11) + F_X(8) = 1 - 0.7483 + 0.2517 = \underline{\underline{0.5034}}$$

Facit: p-värdet är större än $\alpha = 0.05$, vi förkastar **inte** nollhypotesen att $p = 0.5$ och därmed inte heller hypotesen att m_0 är medianen för den underliggande fördelningen. Det är helt möjligt att nollhypotesen stämmer.

Binomialfördelningen, forts.

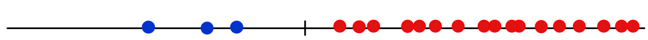
n	$k\backslash p$	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0005	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0036	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0160	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.0510	0.0059
	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.1256	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.2500	0.0577
	7	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.4159	0.1316
	8	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.5956	0.2517
	9	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.7553	0.4119
	10	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.8725	0.5881
	11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9435	0.7483
	12	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9790	0.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

One-Sample Sign test

***p*-värde, fall 1**: H_0 förkastas om x är för litet eller för stort (tvåsidigt test).

$$H_a : m \neq m_0 \quad \Rightarrow \quad p = P\big(X \geq max(x, n-x)\big) + P\big(X \leq min(x, n-x)\big)$$
$$= 2 \cdot P\big(X \leq min(x, n-x)\big)$$
(eftersom $p = 0.5 \rightarrow$ symmetr. fördelning)

Exempel 2: Låt oss anta att vi fick 20 mätvärden ($n = 20$), därav 17 större än den hypotetiska medianen, dvs. $x = 17$.



$$p = P\big(X \geq max(x, n-x)\big) + P\big(X \leq min(x, n-x)\big) \qquad \text{där } X \sim Bin(20, 0.5)$$
$$p = P(X \geq 17) + P(X \leq 3) = 1 - P(X \leq 16) + P(X \leq 3)$$
$$= 1 - F_X(16) + F_X(3) = 1 - 0.9987 + 0.0013 = 0.0026$$

Facit: p-värdet är mindre än $\alpha = 0.05$, vi förkastar H_0, dvs. vi förkastar antagandet att m_0 är medianen för den underliggande fördelningen. Vi accepterar däremot den alternativa hypotesen $m \neq m_0$ (sanna värdet på medianen måste vara något annat än vad hypotesen säger).

Binomialfördelningen, forts.

n	$k\backslash p$	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
20	0	0.3585	0.1216	0.0388	0.0115	0.0032	0.0008	0.0000	0.0000
	1	0.7358	0.3917	0.1756	0.0692	0.0243	0.0076	0.0005	0.0000
	2	0.9245	0.6769	0.4049	0.2061	0.0913	0.0355	0.0036	0.0002
	3	0.9841	0.8670	0.6477	0.4114	0.2252	0.1071	0.0160	0.0013
	4	0.9974	0.9568	0.8298	0.6296	0.4148	0.2375	0.0510	0.0059
	5	0.9997	0.9887	0.9327	0.8042	0.6172	0.4164	0.1256	0.0207
	6	1.0000	0.9976	0.9781	0.9133	0.7858	0.6080	0.2500	0.0577
	7	1.0000	0.9996	0.9941	0.9679	0.8982	0.7723	0.4159	0.1316
	8	1.0000	0.9999	0.9987	0.9900	0.9591	0.8867	0.5956	0.2517
	9	1.0000	1.0000	0.9998	0.9974	0.9861	0.9520	0.7553	0.4119
	10	1.0000	1.0000	1.0000	0.9994	0.9961	0.9829	0.8725	0.5881
	11	1.0000	1.0000	1.0000	0.9999	0.9991	0.9949	0.9435	0.7483
	12	1.0000	1.0000	1.0000	1.0000	0.9998	0.9987	0.9790	0.8684
	13	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9935	0.9423
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9984	0.9793
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9941
	16	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9987
	17	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9998
	18	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
	19	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

One-Sample Sign test

- **Fråga**: vill möss ha en egen spegel ?
- 16 möss med varsin bur som har ett rum med spegel och ett rum utan spegel
- I vilket rum uppehöll sig musen mest?:

mus	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
med		x					x				x					
utan	x		x	x	x	x		x	x	x		x	x	x	x	x

Sherwin, C.M. 2004. Mirrors as potential environmental enrichment for individually housed laboratory mice. Appl. Anim. Behav. Sci. 87: 95-103.

H_0: $p = 0.5$; H_a: $p \neq 0.5$ (tvåsidigt test) $x = 3$ (observation)

$X \sim Bin(16, 0.5)$ ("under H_0")

$$p = P(X \leq 3) + P(X \geq 13) = P(X \leq 3) + 1 - P(X \leq 12) = F_X(3) + 1 - F_X(12)$$
$$= 0.0106 + 1 - 0.9894 = \underline{\underline{0.0212}}$$

Facit: Nollhypotesen ($p = 0.5$) förkastas. Det betyder att musen bryr sig om i vilket rum de är. Mätvärdena avslöjar att musen föredrar ... att **inte** ha någon spegel.

Binomialfördelningen, forts.

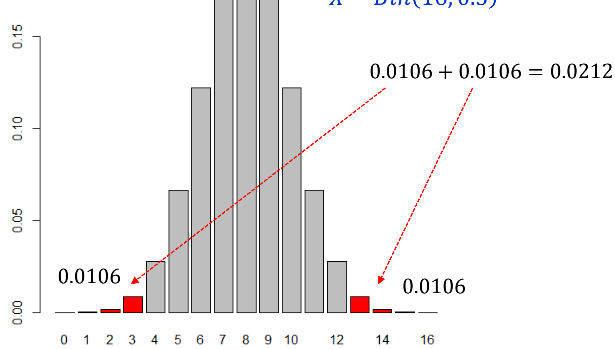
n	$k \backslash p$	0.05	0.1	0.15	0.2	0.25	0.3	0.4	0.5
16	0	0.4401	0.1853	0.0743	0.0281	0.0100	0.0033	0.0003	0.0000
	1	0.8108	0.5147	0.2839	0.1407	0.0635	0.0261	0.0033	0.0003
	2	0.9571	0.7892	0.5614	0.3518	0.1971	0.0994	0.0183	0.0021
	3	0.9930	0.9316	0.7899	0.5981	0.4050	0.2459	0.0651	0.0106
	4	0.9991	0.9830	0.9209	0.7982	0.6302	0.4499	0.1666	0.0384
	5	0.9999	0.9967	0.9765	0.9183	0.8103	0.6598	0.3288	0.1051
	6	1.0000	0.9995	0.9944	0.9733	0.9204	0.8247	0.5272	0.2272
	7	1.0000	0.9999	0.9989	0.9930	0.9729	0.9256	0.7161	0.4018
	8	1.0000	1.0000	0.9998	0.9985	0.9925	0.9743	0.8577	0.5982
	9	1.0000	1.0000	1.0000	0.9998	0.9984	0.9929	0.9417	0.7728
	10	1.0000	1.0000	1.0000	1.0000	0.9997	0.9984	0.9809	0.8949
	11	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997	0.9951	0.9616
	12	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9991	0.9894
	13	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9999	0.9979
	14	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	0.9997
	15	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000

One-Sample Sign Test

p-värdet är sannolikheten av en erhållen observation, eller ännu mer extrema observationer, givet att nollhypotesen är sann.

Sannolikhetsfunktion under H_0

$X \sim Bin(16, 0.5)$



Observation $x = 3$: under H_0 är det mycket osannolikt att en sådan observation kommer till stånd. H_0 förkastas därför.

One-Sample Sign Test

Stort stickprov ($n > 25$)

$$X \sim Bin(n,p)$$

$$X \sim N\left(n \cdot p, \sqrt{n \cdot p \cdot (1-p)}\right) \quad \text{om} \quad V(X) = n \cdot p \cdot (1-p) > 5$$

(se föreläsning **F7**)

$$X \sim Bin(n, 0.5) \quad \text{under } H_0$$

$$X \sim N\left(n/2,\ 1/2\sqrt{n}\right) \quad \text{Normalapproximation, under } H_0$$

$$Z = \frac{X - n/2}{1/2\sqrt{n}} \sim N(0,1)$$

Om nollhypotesen är sann borde Z vara standard-normalfördelad

Variabeln Z är standard-normal fördelad om H_0 stämmer. Under H_0 borde alltså en observation för Z inte ligga för långt ifrån centrumet av täthetsfunktionen för $N(0,1)$. Om observationen däremot ligger i svansen av denna täthetsfunktion förkastar vi nollhypotesen, eftersom observationen är osannolik under H_0. De kritiska områdena är alltså (jämför föreläsning F12):

$$\Omega_{krit} = \{|z| \geq \lambda_{\alpha/2}\} \quad \text{tvåsidigt test } (H_a: p \neq 0.5$$

$$\Omega_{krit} = \{z \geq \lambda_\alpha\} \quad \text{ensidigt test med } H_a: p > 0.5$$

$$\Omega_{krit} = \{z \leq -\lambda_\alpha\} \quad \text{ensidigt test med } H_a: p < 0.5$$

One-Sample Wilcoxon Test

- testar om medianen för en population är lika med ett hypotetiskt värde m_0 (som Sign-testet)
- skillnad till Sign-testet : fördelningen för den underliggande populationen måste vara symmetrisk.

Exempel : se stickprov $\{x_i\}$ i tabellen:

Fråga : Är medianen lika med noll ? ($m_0 = 0$?)

Testvariabel:

- man rangordnar $|x_i|$ (R_i= rang , tabell)
  - (rang 1 = minsta värde)
- om $m_0 = 0$ testas beräknas:
  - rangsumman för positiva x_i (röd i tabellen)
  - rangsumman för negativa x_i (blå i tabellen)

$$W^+ = \sum_i R_i^+ \qquad W^- = \sum_i R_i^-$$

rangsumma för positiva x_i — rangsumma för negativa x_i

$W^+ = 3 + 8 + 5 = 16$

$W^- = 6 + 9 + 4 + 2 + 10 + 11 + 12 + 7 + 1 + 13 = 75$

x_i	$\lvert x_i \rvert$	R_i
-3,1	3,1	6
-6,3	6,3	9
1,2	1,2	3
-2,0	2,0	4
-1,0	1,0	2
-7,2	7,2	10
5,6	5,6	8
2,2	2,2	5
-12,0	12,0	11
-12,3	12,3	12
-5,3	5,3	7
-0,1	0,1	1
-23,4	23,4	13

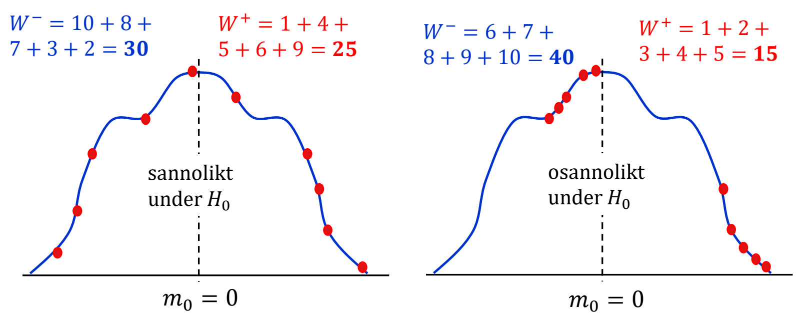
One-Sample Wilcoxon Test

Om medianen av en symmetrisk fördelning är 0

- ... borde W^+ och W^- ungefär vara lika stora

  (se bild till vänster: $W^- = 25$; $W^+ = 30$)
- ... är det däremot osannolikt att t.ex. alla negativa observationer har stora absolutvärden och alla positiva observationer ha små absolutvärden

  (se bild till höger: $W^- = 15$; $W^+ = 40$)



www.matstat.org

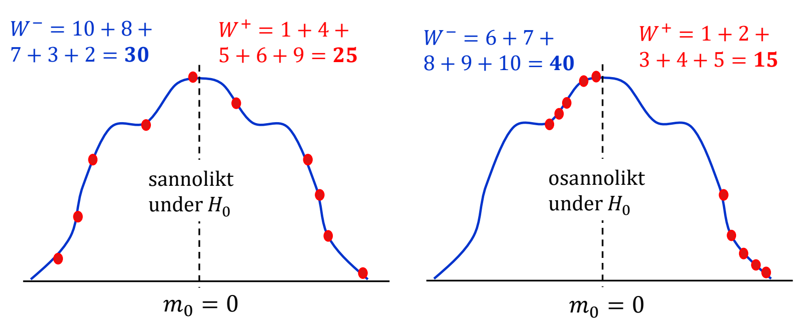
1-S Wilcoxon Test

- H_0 förkastas om skillnaden mellan W^+ och W^- blir för stor.
- ... vilket betyder att en av dem är låg.

⇨ Testvariabel: $W = min(W^+, W^i)$

Kritiskt område: $\Omega_{krit} = \{W < W_0\}$

H_0 förkastas för små värden av testvariabeln



www.matstat.org

One-Sample Wilcoxon test

Tabell över W_0 (kritiska värden)

Critical values **Wilcoxon**

One-sided	Two-sided	$n = 5$	$n = 6$	$n = 7$	$n = 8$	$n = 9$	$n = 10$
$P = .05$	$P = .10$	1	2	4	6	8	11
$P = .025$	$P = .05$		1	2	4	6	8
$P = .01$	$P = .02$			0	2	3	5
$P = .005$	$P = .01$				0	2	3
One-sided	Two-sided	$n = 11$	$n = 12$	$n = 13$	$n = 14$	$n = 15$	$n = 16$
$P = .05$	$P = .10$	14	17	21	26	30	36
$P = .025$	$P = .05$	11	14	17	21	25	30
$P = .01$	$P = .02$	7	10	13	16	20	24
$P = .005$	$P = .01$	5	7	10	13	16	19
One-sided	Two-sided	$n = 17$	$n = 18$	$n = 19$	$n = 20$	$n = 21$	$n = 22$
$P = .05$	$P = .10$	41	47	54	60	68	[illegible]
$P = .025$	$P = .05$	35	40	46	52	59	66
$P = .01$	$P = .02$	28	33	38	43	[illegible]	56
$P = .005$	$P = .01$	23	28	32	37	43	49
One-sided	Two-sided	$n = 23$	$n = 24$	$n = 25$	[illegible]	$n = 27$	$n = 28$
$P = .05$	$P = .10$	83	92	101	110	120	130
$P = .025$	$P = .05$	73	81	90	98	107	117
$P = .01$	$P = .02$	62	69	77	85	93	102
$P = .005$	$P = .01$	55	68	68	76	84	92

beräknas med hjälp av kombinatorik

One-Sample Wilcoxon Test

Stort stickprov ($n > 25$)

$$Z = \frac{W^+ - \mu}{\sigma} \sim N(0,1)$$

standard-normal

$$\mu = \frac{1}{2} \cdot \frac{n \cdot (n+1)}{2}$$

väntevärdet för testvariabel W^+
(under H_0: hälften av den totala rangsumman)

$$\sigma^2 = \frac{n \cdot (n+1) \cdot (2n+1)}{24}$$

varians för testvariabel W^+

$$\Omega_{krit} = \left\{ |z| > \lambda_{\alpha/2} \right\}$$

(för tvåsidigt test)
”större än” - vi testar ju W^+

Wilcoxon-Signed Rank test för parade stickprov

- två parade stickprov (t. ex. ”före/efter”)
- finns det en förskjutning (”location shift”) mellan fördelningarna eller är de identiska (H_0) ?
- Exempel: aluminiumhalt i träd i en förorenad areal i augusti månad och november månad →

Laureysens, I., R. Blust, L. De Temmerman, C. Lemmens and R. Ceulemans. 2004. Clonal variation in heavy metal accumulation and biomass production in a poplar coppice culture. I. Seasonal variation in leaf, wood and bark concentrations. Environ. Pollution 131: 485-494.

www.matstat.org

Wilcoxon-Signed Rank Test för parade stickprov

- de absoluta differenserna ($|\Delta_i|$) rangordnas → R_i (rang 1 = minsta)
- beräkna W^+ och W^- igen:

$$W^+ = \sum_i R_i^+ \quad W^- = \sum_i R_i^-$$

$W^+ = 3 + 8 + 5 = \mathbf{16}$

$W^- = 6+9+4+2+10+11+12+7+1+13 = \mathbf{75}$

www.matstat.org

Aug	Nov	Δ_i	$\|\Delta_i\|$	R_i
8,1	11,2	-3,1	3,1	6
10,0	16,3	-6,3	6,3	9
16,5	15,3	1,2	1,2	3
13,6	15,6	-2,0	2,0	4
9,5	10,5	-1,0	1,0	2
8,3	15,5	-7,2	7,2	10
18,3	12,7	5,6	5,6	8
13,3	11,1	2,2	2,2	5
7,9	19,9	-12,0	12,0	11
8,1	20,4	-12,3	12,3	12
8,9	14,2	-5,3	5,3	7
12,6	12,7	-0,1	0,1	1
13,4	36,8	-23,4	23,4	13

Wilcoxon-Signed Rank Test för parade stickprov

- Om båda fördelningarna var identiska borde ungefär hälften av de parvisa differenserna vara positiva och den andra hälften vara negativa.
- Dessutom borde positiva och negativa differenser av samma storlek vara lika sannolika.
- Under H_0 förväntar vi oss alltså att rangsumman för de positiva differenserna och rangsumman för de negativa differenserna är ungefär lika stora.
- Vi genomför alltså ett 1-S Wilcoxon test för de parvisa differenserna :

$$W^+ = \sum_i R_i^+ \qquad W^- = \sum_i R_i^-$$

rangsumman för positiva Δ_i rangsumman för negativa Δ_i

$$\left.\begin{array}{l} W^+ = 3 + 8 + 5 = 16 \\ W^- = 6+9+4+2+10+11+12+7+1+13=75 \end{array}\right\} \text{se tabell ovan}$$

Testvariabel: $W = min(W^+, W^i)$

H_0 förkastas för små värden av testvariabeln

Kritiskt område: $\Omega_{krit} = \{W < W_0\}$

www.matstat.org

Mann-Whitney U-test

Syftet:

- H_0 : två oberoende (icke-parade) stickprov kommer från identiska populationer
- H_a : det finns en förskjutning mellan populationerna (”location shift”)
- motsvarar Student’s t-test för icke-normalfördelade populationer

<u>Exempel:</u> två stickprov A och B:

- $x_A = \{25, 26, 27, 31\}$ och $x_B = \{28, 29, 32, 35\}$ (i praktiken mycket för små)

Testvariabel (U):

- ordnar alla värden enligt deras storlek (minsta värde först)
- räkna: hur många värden av A kommer före varje B-värde ? → $\boldsymbol{U_B}$
- räkna: hur många värden av B kommer före varje A-värde ? → $\boldsymbol{U_A}$

rangordnat värde	25	26	27	28	29	31	32	35
från stickprov:	A	A	A	B	B	A	B	B
bidrag till U_B				3	3		4	4

3 värden från stickprov A kommer före denna B

4 värden från stickprov A kommer före denna B

www.matstat.org

Testvariabel, Mann-Whitney U-test

- räkna: hur många värden av A kommer före varje B-värde ? → $\boldsymbol{U_B}$
- räkna: hur många värden av B kommer före varje A-värde ? → $\boldsymbol{U_A}$

rangordnat värde	25	26	27	28	29	31	32	35
från stickprov:	A	A	A	B	B	A	B	B
bidrag till U_B				3	3		4	4

3 värden från stickprov A kommer före denna B

4 värden från stickprov A kommer före denna B

$$\boldsymbol{U_B = 3 + 3 + 4 + 4 = 14}$$

rangordnat värde	25	26	27	28	29	31	32	35
från stickprov:	A	A	A	B	B	A	B	B
bidrag till U_A	0	0	0			2		

0 värden från stickprov B kommer före denna A

2 värden från stickprov B kommer före detta A

$$\boldsymbol{U_A = 0 + 0 + 0 + 2 = 2}$$

Mann-Whitney U-test

Testvariabel (U):

$$\boxed{U = min(U_A, U_B)}$$

Kritiskt område:

- mycket små värden för testvariabeln U tyder på en skillnad mellan fördelningarna för populationerna A och B
- det kritiska värdet (U_0) fås ur en tabell

$$\boxed{\Omega_{krit} = \{U < U_0\}}$$

Förutsättningar; styrka (power):

- oberoende slumpmässiga stickprov
- (lite) mindre styrka jämfört med Student's t-test ifall populationerna är normalfördelade
- större styrka än Student's t-test för många andra fördelningar

www.matstat.org

Mann-Whitney U-test

U och W (Wilcoxon) hänger ihop : n_1, n_2 stickprovsstorlekar

värde	25	26	27	28	29	31	32	35
stickprov	A	A	A	B	B	A	B	B
bidrag till U_B				3	3		4	4
bidrag till U_A	0	0	0			2		
rang (Wilcoxon)	1	2	3	4	5	6	7	8

$$\left.\begin{aligned} W_A &= 1+2+3+6 = 12 \\ W_B &= 4+5+7+8 = 24 \end{aligned}\right\} \text{ Wilcoxons rangsumma}$$

$$\left.\begin{aligned} U_A &= n_1 \cdot n_2 + \frac{n_1 \cdot (n_1+1)}{2} - W_A = 16 + \frac{20}{2} - 12 = 14 \\ U_B &= n_1 \cdot n_2 + \frac{n_2 \cdot (n_2+1)}{2} - W_B = 16 + \frac{20}{2} - 24 = 2 \end{aligned}\right\} \text{ sammanhang mellan } U \text{ och } W$$

$$U_A + U_B = n_1 \cdot n_2 = 16$$

Dessa formler används ofta i praktiken (beräkning av W enklare).

www.matstat.org

Mann-Whitney U-test

Stort stickprov ($n > 25$)

$$Z = \frac{U - \mu_U}{\sigma_U} \sim N(0,1)$$ ... är standard-normal

$$\mu_U = \frac{n_1 \cdot n_2}{2}$$ väntevärdet för testvariabel U

$$\sigma_U = \sqrt{\frac{n_1 \cdot n_2 \cdot (n_1 + n_2 + 1)}{12}}$$ standardavvikelse för testvariabel U

$$\Omega_{krit} = \left\{ |z| > \lambda_{\alpha/2} \right\}$$ kritiskt område (för tvåsidigt test)

Alternativa namn för Mann-Whitney U-testet:

- Mann–Whitney–Wilcoxon test
- Wilcoxon rank-sum test
- Wilcoxon–Mann–Whitney test

www.matstat.org

Kruskal-Wallis Test

- används istället för One-Way ANOVA ifall populationerna inte är normalfördelade
- k: antalet populationer (”grupper”, ”treatments”)
- H_0: alla k populationer har samma fördelning
- H_a: minst en population har en förskjuten fördelning (”location shift”)

Testvariabel:

- alla $n = n_1 + n_2 + \cdots + n_k$ observationer rangordnas
  - n_i = antalet värden i stickprovet för grupp i

R_i: summan av alla ranger för stickprov i

$\bar{R}_i = \frac{R_i}{n_i}$ medelvärdet av alla ranger för stickprov i

$\bar{R} = \frac{1}{n} \cdot \frac{n \cdot (n+1)}{2} = \frac{n+1}{2}$ medelvärdet av alla ranger (medelvärdet av de första n naturliga tal)

$\frac{n \cdot (n+1)}{2}$ = summan av alla ranger = summan av de första n naturliga talen

www.matstat.org

Kruskal-Wallis Test

Definition: $V = \sum_{i=1}^{k} n_i \cdot \left(\bar{R}_i - \bar{R}\right)^2$

V motsvarar kvadratsumman SS_{tr} för ANOVA (se föreläsning **F13**):

- om H_0 var sann borde alla $\bar{R}_i$ (och $\bar{R}$) vara ungefär lika stora
- $\rightarrow V$ måste vara liten under H_0

Testvariabel: (liten under H_0) $H = \frac{12 \cdot V}{n \cdot (n+1)} = \frac{12}{n \cdot (n+1)} \cdot \sum_{i=1}^{k} \frac{R_i^2}{n_i} - 3 \cdot (n+1)$

Kritiskt område:

- Om H_0 gäller och dessutom alla n_i är tillräckligt stora ($n_i \geq 5$) är testvariabeln H ungefär χ^2- fördelad.
- H_0 förkastas om H (alltså även V) är stor – detta betyder ju att gruppmedelvärderna $\bar{R}_i$ avviker mycket från det gemensamma medelvärdet $\bar{R}$.

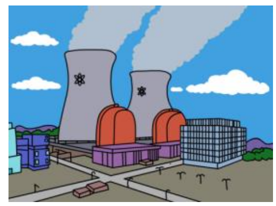
$$\Omega_{krit} = \{\, H > \chi^2_\alpha(f) \,\} \qquad f = k - 1$$

www.matstat.org

Kruskal-Wallis Test

Experiment: Vi vill veta om temperaturförhöjningen av havsvattnet i närheten av ett kärnkraftverk har ett inflytande på fiskarnas vikt:

Vikten av fisk			
38˚F	**42˚F**	**46˚F**	**50˚F**
22	15	14	17
24	21	28	18
16	26	21	13
18	16	19	20
19	25	24	21
	17	23	



- 4 grupper
- $n_i = (5, 6, 6, 5)$
- $n = 22$

www.matstat.org

Kruskal-Wallis Test

kruskal_wallis.R (se webbsida)

```
x_38 = c(22, 24, 16, 18, 19)
x_42 = c(15, 21, 26, 16, 25, 17)
x_46 = c(14, 28, 21, 19, 24, 23)
x_50 = c(17, 18, 13, 20, 21)
weight = c(x_38, x_42, x_46, x_50)
temp = c( rep(38, length(x_38)), rep(42, length(x_42)),
          rep(46, length(x_46)), rep(50, length(x_50)))
fish = data.frame(weight = weight, temp = temp)
boxplot(weight ~ temp, data = fish)
kruskal.test(weight ~ temp, data = fish) # Kruskal-Wallis rank sum Test
# Kruskal-Wallis chi-squared = 2.0404, df = 3, p-value = 0.5641
```

www.matstat.org