KTH Teknikvetenskap
Harald Lang

Formelsamling och Tabeller i Statistik och Sannolikhetsteori

(15/11-10)

Datareducering

- Om $x_1, \ldots, x_n$ är ett stickprov ur en population så definieras *medelvärdet* $\bar{x}$
  $\bar{x} = \frac{1}{n}\sum_{k=1}^{n} x_k$
  och *standardavvikelsen* s
  $s = \sqrt{\frac{1}{n-1}\sum_{k=1}^{n}(x_k - \bar{x})^2}$. På miniräknaren kan s ha beteckningen $x\sigma_{n-1}$, $\sigma_{x,n-1}$, s_x eller något snarlikt.

Kombinatorik

- Antalet sätt att välja ut r objekt bland n stycken *utan hänsyn till ordning* är
  $\binom{n}{r} = \frac{n(n-1)\cdots(n-r+1)}{r!}$. På miniräknaren skrivs detta som ${}_nC_r$.
  $\binom{n}{0} = \binom{n}{n} = 1$, $\binom{n}{r} = \binom{n}{n-r}$.

- Antalet sätt att välja ut r objekt bland n stycken *med hänsyn till ordning* är
  ${}_nP_r = n(n-1)\cdots(n-r+1)$.

- Antalet sätt att dra r objekt bland n stycken *med återläggning och utan hänsyn till ordning* är $\binom{n+r-1}{r}$.

Sannolikheter

P(A^c n B^c) = 1 - P(A U B)

- $P(A \cup B) = P(A) + P(B)$ om händelserna A och B är *uteslutande* (oförenliga, disjunkta).
  I allmänhet gäller
  $P(A \cup B) = P(A) + P(B) - P(A \cap B)$.

  3 mängder, plussa på A n B n C

- $P(A \mid B) = \frac{P(A \cap B)}{P(B)}$ är sannolikheten för A *betingat händelsen* B.
  Händelserna A *och* B *är oberoende* om $P(A \cap B) = P(A)P(B)$, vilket är ekvivalent med $P(B \mid A) = P(B)$ och naturligtvis även ekvivalent med $P(A \mid B) = P(A)$.

- Generellt gäller
  $P(A \cap B) = P(A)P(B \mid A) = P(B)P(A \mid B)$.

- *Lagen om total sannolikhet* är identiteten
  $P(A) = P(B)P(A \mid B) + P(B^*)P(A \mid B^*)$.
  (B^* betecknar händelsen att B *inte* inträffar.)
  Mer generellt: Om händelserna $H_1, \dots, H_n$ är sådana att precis en måste inträffa, så gäller att
  $P(A) = P(H_1)P(A \mid H_1) + \cdots + P(H_n)P(A \mid H_n)$.

- *Bayes Regel* är BETINGAD SANNOLIKHET
  $P(B \mid A) = \dfrac{P(B)P(A \mid B)}{P(A)}$ där nämnaren kan beräknas enligt ovan.

Fördelningar och Stokastiska Variabler

Stora bokstäver $X,\ Y,\ Z$ betecknar stokastiska variabler, små bokstäver $x,\ y,\ z$ betecknar värden.

- *Väntevärdet* (för en diskret variabel) $\mathrm{E}\,[X]$ definieras som
  $\mathrm{E}\,[X] = \mu = \sum_n x_n P(X = x_n)$
  där $x_1, x_2,\ \dots$ är en uppräkning av alla värden som X kan anta.

  E(X) = = x * p(x)

- *Standardavvikelsen* (för en diskret variabel) $\mathrm{SD}\,[X]$ definieras som
  $\mathrm{SD}\,[X] = \sigma = \sqrt{\sum_n (x_n - \mu)^2 P(X = x_n)}$.

  SD(X) = (E(X^2) - ^2)

- *Kovariansen* mellan X och Y är $\mathrm{Cov}[X, Y] = \mathrm{E}\,[XY] - \mathrm{E}\,[X]\,\mathrm{E}\,[Y]$.
  Variansen för X är $\mathrm{Var}[X] = \mathrm{Cov}[X, X]$; $\mathrm{SD}[X] = \sqrt{\mathrm{Var}[X]}$.

  Var(X) = E(X^2) - ^2

- *Korrelationskoefficienten* $\mathrm{Corr}[X, Y] = \rho = \dfrac{\mathrm{Cov}[X, Y]}{\mathrm{SD}[X]\,\mathrm{SD}[Y]}$.

- **Bernoulli-fördelning.** I är *Bernoulli-fördelad* (p) om I bara kan anta värdena 0 och 1 :
  $$\begin{cases} P(X = 1) = p\,, \\ P(X = 0) = 1 - p\,. \end{cases}$$
  $\mathrm{E}\,[I] = p\,;\ \ \mathrm{SD}\,[I] = \sqrt{p(1 - p)}$.

- **Binomialfördelning.** X är *binomialfördelad* (n, p), skrivs $\mathrm{Bin}(n, p)$, om
  $X = I_1 + \cdots + I_n$
  där $I_1, \dots, I_n$ är *oberoende* Bernoulli (p)-variabler. Typiskt exempel är att ett experiment utförs n gånger, och att experimenten ”lyckas” med sannolikheten p varje gång, oberoende av varandra. Antalet lyckade försök blir då $\mathrm{Bin}(n, p)$.
  $$P(X = r) = \binom{n}{r} p^r\,(1 - p)^{n - r}, \quad r = 0, \dots, n.$$
  $\mathrm{E}\,[X] = np$
  $\mathrm{SD}\,[X] = \sqrt{np\,(1 - p)}$.

- **Hypergeometrisk fördelning.** Typexempel: Man väljer ut n objekt utan hänsyn till ordning bland N stycken, utan återläggning. Antag att av de N objekten a stycken är defekta, medan resten $b = N - a$ inte är defekta. Då är sannolikheten att man får precis r defekta objekt

Multinomialfördelning

$\boxed{\text{P = (n! / n1! n2! n3!) * (p^n1 * p^n2 * p^n3)}}$

$$\frac{\binom{a}{r}\binom{b}{n-r}}{\binom{N}{n}}.$$

Om X betecknar antalet defekta enheter man valt ut är alltså uttrycket ovan $P(X = r)$. Det gäller att

$$\mathrm{E}[X] = \frac{na}{N}$$

$$\mathrm{SD}[X] = \sqrt{\frac{N-n}{N-1}\frac{nab}{N^2}}.$$

- **Poissonfördelning.** Antag att händelser inträffar slumpmässigt oberoende av varandra med en viss intensitet λ. Intensiteten λ är genomsnittliga antalet inträffade händelser under observationsperioden. Om X betecknar det faktiska antalet inträffade händelser vid en observation, är X *Poissonfördelad,* med intensitet λ, beteckning $\mathrm{Po}(\lambda)$. Det gäller att

$$P(X = r) = \frac{\lambda^r}{r!}e^{-\lambda},\ r = 0, 1, 2, \ldots, \quad \mathrm{E}[X] = \lambda, \quad \mathrm{SD}[X] = \sqrt{\lambda}.$$

- **"För-första-gången"-fördelningen.** Om en händelse inträffar med sannolikheten p och X är antalet oberoende försök tills dess händelsen inträffar, är X "för-första-gången"-fördelad, $X \in \mathrm{Ffg}(p)$.

$$P(X = r) = p(1-p)^{r-1},\ r = 1, 2, \ldots, \quad \mathrm{E}[X] = \frac{1}{p}, \quad \mathrm{SD}[X] = \sqrt{\frac{1}{p^2} - \frac{1}{p}}.$$

- **Normalfördelning.** Täthetsfunktion $f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$. $\mathrm{E}[X] = 0$, $\mathrm{SD}[X] = 1$.

Detta är *standard-Normalfördelningen.* Om X är standard-Normalfördelad skriver vi $X \in N(0,1)$ (nollan indikerar väntevärdet, ettan standardavvikelsen.)

<u>Om $X \in \mathrm{Bin}(n,p)$</u> och $np(1-p) > 10$ gäller att $Z = \frac{X-\mu}{\sigma} \approx N(0,1)$, dvs. Z är approximativt (standard-)Normalfördelad. Här är $\mu = np$ och $\sigma = \sqrt{np(1-p)}$. Om man skall få en bra approximation bör man göra en "kontinuitets-korrektion:" $P(k \le X \le m) \approx P\left(\frac{(k-0.5)-\mu}{\sigma} \le Z \le \frac{(m+0.5)-\mu}{\sigma}\right)$.

<u>Om $X \in \mathrm{Po}(\lambda)$</u> och $\lambda > 15$ gäller att $Z = \frac{X-\lambda}{\sqrt{\lambda}} \approx N(0,1)$, dvs. Z är approximativt (standard-)Normalfördelad. Om man skall få en bra approximation bör man göra en "kontinuitets-korrektion:" $P(k \le X \le m) \approx P\left(\frac{(k-0.5)-\lambda}{\sqrt{\lambda}} \le Z \le \frac{(m+0.5)-\lambda}{\sqrt{\lambda}}\right)$.

- **Exponentialfördelning.** X är exponentialfördelad om den har täthetsfunktionen

  $$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{för } x \geq 0 \\ 0 & \text{för } x < 0. \end{cases}$$

  Vi skriver $X \in \text{Exp}(\lambda)$. $\mathrm{E}[X] = \dfrac{1}{\lambda}$, $\mathrm{SD}[X] = \dfrac{1}{\lambda}$.

- **Medelvärden av stickprov.** Om $X_1, \ldots, X_n$ är *oberoende* observationer ur samma fördelning med $\mathrm{E}[X] = \mu$ och $\mathrm{SD}[X] = \sigma$, och $\overline{X}$ betecknar deras medelvärde $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$, så gäller att $\mathrm{E}[\overline{X}] = \mu$ och $\mathrm{SD}[\overline{X}] = \dfrac{\sigma}{\sqrt{n}}$.

- **Centrala Gränsvärdessatsen (CGS).** Om $X_1, \ldots, X_n$ är *oberoende* observationer ur samma fördelning med $\mathrm{E}[X] = \mu$ och $\mathrm{SD}[X] = \sigma$, och $\overline{X}$ betecknar deras medelvärde $\overline{X} = \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} X_k$, så gäller att $Z = \dfrac{\overline{X} - \mu}{\sigma/\sqrt{n}} \approx N(0,1)$. Förutsättningen är att n är stort, men tyvärr finns ingen bra tumregel för hur stort n skall vara.

Konfidensintervall

- **Konfidensintervall för väntevärdet.** Om $x_1, \ldots, x_n$ är oberoende observationer ur en fördelning som är (approximativt) normalfördelad, så är

  $$\mu = \bar{x} \pm t_{\alpha/2} \cdot \frac{s}{\sqrt{n}} \qquad \text{(symmetriskt)}$$

  $$\mu \leq \bar{x} + t_{\alpha} \cdot \frac{s}{\sqrt{n}} \qquad \text{(enkelsidigt)}$$

  $$\mu \geq \bar{x} - t_{\alpha} \cdot \frac{s}{\sqrt{n}} \qquad \text{(enkelsidigt)}$$

  konfidensintervall för väntevärdet μ med konfidensgraden $1 - \alpha$. Här är t_a a-kvantilen för t-fördelningen med $f = n - 1$ frihetsgrader.

  Om standardavvikelsen σ är känd använder man σ i stället för s och sätter $f = \infty$.

  Om fördelningen inte är (approximativt) normalfördelad men $\bar{x}$ kan antas komma från en approximativ normalfördelning (pga. CGS, t.ex.) kan man använda formlerna med $f = \infty$ och någon punktskattning av σ för s.

- **Konfidensintervall för Poisson-intensitet**. Om vi har n observationer vars summa är k av en $\text{Po}(\lambda)$-variabel, så är ett konfidensintervall med approximativa felrisken α för λ

  $$\lambda = \frac{k+2}{n} \pm z_{\alpha/2}\frac{\sqrt{k+1}}{n} \qquad \text{(symmetriskt)}$$

  $$\lambda \leq \frac{k+2}{n} + z_{\alpha}\frac{\sqrt{k+1}}{n} \qquad \text{(enkelsidigt)}$$

  $$\lambda \geq \frac{k+2}{n} - z_{\alpha}\frac{\sqrt{k+1}}{n} \qquad \text{(enkelsidigt.)}$$

Här är z_a a-kvantilen för normalfördelningen ($f = \infty$ i t-fördelningen). För att approximationen skall vara någorunda bra bör $k \geq 10$.

- **Konfidensintervall för andel** (eller sannolikhet). Om k av n slumpvis utvalda objekt ur en "oändlig" population har en egenskap E, så är konfidensintervall med approximativ felrisk α för p, dvs. andelen objekt i hela populationen som har egenskapen E med approximativa felrisken α för λ

  $$\hat{p} = \frac{k+2}{n+4}$$

  $$p = \hat{p} \pm z_{\alpha/2} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n+4}} \qquad \text{(symmetriskt)}$$

  $$p \leq \hat{p} + z_{\alpha} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n+4}} \qquad \text{(enkelsidigt)}$$

  $$p \geq \hat{p} - z_{\alpha} \cdot \sqrt{\frac{\hat{p}(1-\hat{p})}{n+4}} \qquad \text{(enkelsidigt.)}$$

  Här är z_a a-kvantilen för normalfördelningen ($f = \infty$ i t-fördelningen). För att approximationen skall vara någorunda bra bör $k(n-k) \geq 5n$.

- **Konfidensintervall för skillnad i väntevärden.** Om $x_1, \ldots, x_n$ och $y_1, \ldots, y_m$ är oberoende observationer ur fördelningar med väntevärde μ_x respektive μ_y så är ett approximativt konfidensintervall för differensen $\mu_x - \mu_y$

  $$\mu_x - \mu_y = \bar{x} - \bar{y} \pm z_{\alpha/2} \cdot \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} \qquad \text{(symmetriskt)}$$

  $$\mu_x - \mu_y \leq \bar{x} - \bar{y} + z_{\alpha} \cdot \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} \qquad \text{(enkelsidigt)}$$

  $$\mu_x - \mu_y \geq \bar{x} - \bar{y} - z_{\alpha} \cdot \sqrt{\frac{s_x^2}{n} + \frac{s_y^2}{m}} \qquad \text{(enkelsidigt)}$$

  Kravet är att $\bar{x}$ och $\bar{y}$ kan antas komma från approximativa normalfördelningar (pga. CGS, t.ex.). s_x och s_y kan ersättas med andra punktskattningar av standardavvikelserna för respektive fördelning. Här är z_a a-kvantilen för normalfördelningen ($f = \infty$ i t-fördelningen); beteckningar enligt "Datareducering".

Hypotesprövning

- **Hypotesprovning av median (medelvärde).** Om fördelningen är approximativt normalfördelad, använd relevant *konfidensintervall* för väntevärdet, annars *teckensatta rangtestet.*

- **Hypotesprövning av lika medianer (medelvärden, fördelning):** Rangsummetestet (alternativt: konfindesintervall.)

- **Hypotesprövning av proportion:** χ^2-test. **Jämförelse av proportioner:** kontingenstabell. För ensidigt test på nivån α, $\mathrm{H}_0 : p \leq p_0$, gör χ^2-testet på nivån 2α, men förkasta bara om $\hat{p} > p_0$. På samma sätt om $\mathrm{H}_0 : p_1 \leq p_2$, gör testet på nivån 2α, men förkasta bara om $\hat{p}_1 > \hat{p}_2$.

- **Hypotesprövning av Poisson-intensitet:** konfidensintervall. **Jämförelse av Poisson-intensiteter:** χ^2-test med $f = c - 1$.

χ^2-test

- **Kontingenstabell.**

$$\begin{array}{cccc|c} x_{11} & x_{12} & \cdots & x_{1c} & n_1 \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{r1} & x_{r2} & \cdots & x_{rc} & n_r \\ \hline m_1 & m_2 & \cdots & m_c & N \end{array}$$

$n_1, \ldots, n_r$ är radsummor, $m_1, \ldots, m_c$ är kolonnsummor, N är totalsumman. Antalet frihetsgrader är $f = (r-1)(c-1)$.

$$Q = \sum_{i=1}^{r} \sum_{j=1}^{c} \frac{(x_{ij} - n_i m_j / N)^2}{n_i m_j / N}.$$

- **Test för fördelning, proportion (sannolikhet).** Låt $A_1, \ldots, A_c$ vara uteslutande händelser där någon måste inträffa. Låt $p_1, \ldots, p_c$ vara hypotesen om sannolikheter för dessa händelser: $P(A_k) = p_k$, där $\sum_{k=1}^{c} p_k = 1$. Man har n observationer där frekvensen för händelse A_k är x_k. Testvariabel

$$Q = \sum_{k=1}^{c} \frac{(x_k - np_k)^2}{np_k}.$$

Denna är approximativt χ^2-fördelad med $f = c - 1$ frihetsgrader.
Om man skattat k parametrar för p_k-na, så blir antalet frihetsgrader $f = c - k - 1$.

Icke-Parametriska Test

- **Teckentestet.** Använd binomialfördelningen.

- **Wilcoxons Teckensatta Rangtest.** Test av median $\tilde{\mu}$ på nivå α. Låt t^+ vara rangsumman av skillnader för observationer $> \tilde{\mu}_0$ och t^- rangsumman av skillnader för observationer $< \tilde{\mu}_0$.

  Förkasta $\mathrm{H}_0 : \tilde{\mu} \leq \tilde{\mu}_0$ om $t^- \leq T_{2\alpha}$; förkasta $\mathrm{H}_0 : \tilde{\mu} \geq \tilde{\mu}_0$ om $t^+ \leq T_{2\alpha}$; förkasta $\mathrm{H}_0 : \tilde{\mu} = \tilde{\mu}_0$ om $\min\{t^+, t^-\} \leq T_\alpha$.

- **Wilcoxons Rangsummetest.** Test av skillnad i median på nivå α. Egentligen testar man hypotesen $P(X > Y) = 0.5$. Låt w_1 och w_2 vara rangsummorna av observationerna för serie 1 respektive 2 i det sammanslagna materialet.

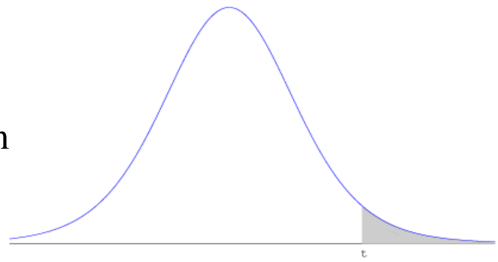
  Definiera $u_i = w_i - n_i(n_i + 1)/2,\ i = 1, 2$ (n_i är antalet observationer i serie i.) Förkasta $\mathrm{H}_0 : \tilde{\mu}_1 \leq \tilde{\mu}_2$ om $u_2 \leq U_{2\alpha}$; förkasta $\mathrm{H}_0 : \tilde{\mu}_1 \geq \tilde{\mu}_2$ om $u_1 \leq U_{2\alpha}$; förkasta $\mathrm{H}_0 : \tilde{\mu}_1 = \tilde{\mu}_2$ om $\min\{u_1, u_2\} \leq U_\alpha$.

<u>Normalfördelningen.</u>
Tabellen anger arean under kurvan mellan 0 och z (alternativt mellan $-z$ och 0).

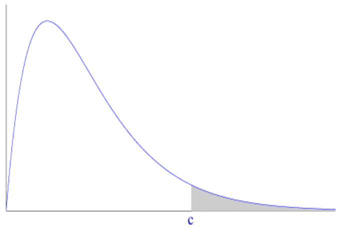
	0	1	2	3	4	5	6	7	8	9
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993

För kvantiler: se t-tabellen med $f=\infty$

Kvantiler för
t-fördelningen.
Tabellen visar värdet på *t* som ger arean i tabellhuvudet till höger om t; ***f***=frihetsgrader.



f	0,1	0,05	0,025	0,01	0,005	0,001	0,0005
1	3,0777	6,3138	12,7062	31,8205	63,6567	318,3088	636,6192
2	1,8856	2,9200	4,3027	6,9646	9,9248	22,3271	31,5991
3	1,6377	2,3534	3,1824	4,5407	5,8409	10,2145	12,9240
4	1,5332	2,1318	2,7764	3,7469	4,6041	7,1732	8,6103
5	1,4759	2,0150	2,5706	3,3649	4,0321	5,8934	6,8688
6	1,4398	1,9432	2,4469	3,1427	3,7074	5,2076	5,9588
7	1,4149	1,8946	2,3646	2,9980	3,4995	4,7853	5,4079
8	1,3968	1,8595	2,3060	2,8965	3,3554	4,5008	5,0413
9	1,3830	1,8331	2,2622	2,8214	3,2498	4,2968	4,7809
10	1,3722	1,8125	2,2281	2,7638	3,1693	4,1437	4,5869
11	1,3634	1,7959	2,2010	2,7181	3,1058	4,0247	4,4370
12	1,3562	1,7823	2,1788	2,6810	3,0545	3,9296	4,3178
13	1,3502	1,7709	2,1604	2,6503	3,0123	3,8520	4,2208
14	1,3450	1,7613	2,1448	2,6245	2,9768	3,7874	4,1405
15	1,3406	1,7531	2,1314	2,6025	2,9467	3,7328	4,0728
16	1,3368	1,7459	2,1199	2,5835	2,9208	3,6862	4,0150
17	1,3334	1,7396	2,1098	2,5669	2,8982	3,6458	3,9651
18	1,3304	1,7341	2,1009	2,5524	2,8784	3,6105	3,9216
19	1,3277	1,7291	2,0930	2,5395	2,8609	3,5794	3,8834
20	1,3253	1,7247	2,0860	2,5280	2,8453	3,5518	3,8495
21	1,3232	1,7207	2,0796	2,5176	2,8314	3,5272	3,8193
22	1,3212	1,7171	2,0739	2,5083	2,8188	3,5050	3,7921
23	1,3195	1,7139	2,0687	2,4999	2,8073	3,4850	3,7676
24	1,3178	1,7109	2,0639	2,4922	2,7969	3,4668	3,7454
25	1,3163	1,7081	2,0595	2,4851	2,7874	3,4502	3,7251
26	1,3150	1,7056	2,0555	2,4786	2,7787	3,4350	3,7066
27	1,3137	1,7033	2,0518	2,4727	2,7707	3,4210	3,6896
28	1,3125	1,7011	2,0484	2,4671	2,7633	3,4082	3,6739
29	1,3114	1,6991	2,0452	2,4620	2,7564	3,3962	3,6594
30	1,3104	1,6973	2,0423	2,4573	2,7500	3,3852	3,6460
40	1,3031	1,6839	2,0211	2,4233	2,7045	3,3069	3,5510
60	1,2958	1,6706	2,0003	2,3901	2,6603	3,2317	3,4602
120	1,2886	1,6577	1,9799	2,3578	2,6174	3,1595	3,3735
∞	1,2816	1,6449	1,9600	2,3263	2,5758	3,0902	3,2905



Sista raden:
För $f > 1000$ kan man använda formeln

$[1-2/(9f) + z\sqrt{2/(9f)}\,]^3$

där z är talet i sista raden.

Kvantiler för chi2/*f*-fördelningen.

Tabellen visar värdet c på X/f som ger arean i tabellhuvudet under kurvan till höger om c. f=frihetsgrader.

f	0,1	0,05	0,02	0,01	0,005
1	2,7055	3,8415	5,4119	6,6349	7,8794
2	2,3026	2,9957	3,9120	4,6052	5,2983
3	2,0838	2,6049	3,2791	3,7816	4,2794
4	1,9449	2,3719	2,9170	3,3192	3,7150
5	1,8473	2,2141	2,6776	3,0173	3,3499
6	1,7741	2,0986	2,5055	2,8020	3,0913
7	1,7167	2,0096	2,3746	2,6393	2,8968
8	1,6702	1,9384	2,2710	2,5113	2,7444
9	1,6315	1,8799	2,1866	2,4073	2,6210
10	1,5987	1,8307	2,1161	2,3209	2,5188
11	1,5705	1,7887	2,0562	2,2477	2,4324
12	1,5458	1,7522	2,0045	2,1847	2,3583
13	1,5240	1,7202	1,9593	2,1299	2,2938
14	1,5046	1,6918	1,9195	2,0815	2,2371
15	1,4871	1,6664	1,8840	2,0385	2,1868
16	1,4714	1,6435	1,8521	2,0000	2,1417
17	1,4570	1,6228	1,8232	1,9652	2,1011
18	1,4439	1,6039	1,7970	1,9336	2,0642
19	1,4318	1,5865	1,7730	1,9048	2,0306
20	1,4206	1,5705	1,7510	1,8783	1,9998
21	1,4102	1,5557	1,7306	1,8539	1,9715
22	1,4006	1,5420	1,7118	1,8313	1,9453
23	1,3916	1,5292	1,6943	1,8104	1,9209
24	1,3832	1,5173	1,6779	1,7908	1,8983
25	1,3753	1,5061	1,6626	1,7726	1,8771
26	1,3678	1,4956	1,6483	1,7554	1,8573
27	1,3608	1,4857	1,6348	1,7394	1,8387
28	1,3541	1,4763	1,6221	1,7242	1,8212
29	1,3478	1,4675	1,6101	1,7099	1,8047
30	1,3419	1,4591	1,5987	1,6964	1,7891
35	1,3160	1,4229	1,5498	1,6383	1,7221
40	1,2951	1,3940	1,5109	1,5923	1,6692
45	1,2779	1,3701	1,4790	1,5546	1,6259
50	1,2633	1,3501	1,4523	1,5231	1,5898
60	1,2399	1,3180	1,4097	1,4730	1,5325
70	1,2218	1,2933	1,3770	1,4346	1,4888
80	1,2072	1,2735	1,3509	1,4041	1,4540
100	1,1850	1,2434	1,3114	1,3581	1,4017
120	1,1686	1,2214	1,2827	1,3246	1,3637
140	1,1559	1,2044	1,2605	1,2989	1,3346
160	1,1457	1,1907	1,2428	1,2783	1,3114
180	1,1372	1,1795	1,2282	1,2614	1,2923
200	1,1301	1,1700	1,2159	1,2472	1,2763
250	1,1162	1,1515	1,1922	1,2198	1,2454
300	1,10596	1,13798	1,17475	1,19969	1,22281
400	1,09162	1,11908	1,15053	1,17181	1,19152
500	1,08186	1,10625	1,13414	1,15299	1,17041
750	1,06672	1,08643	1,10889	1,12404	1,13802
1000	1,05772	1,07468	1,09398	1,10697	1,11895
z	1,28155	1,64485	2,05375	2,32635	2,57583

Kritiska värden för Wilcoxons test med teckensatta ranger.

Tabellen ger de kritiska värdena T_α, dvs tal sådana att sannolikheten att $t \leq T_\alpha$ är så nära α som möjligt.

α	0.10	0.05	0.02	0.01
5	1			
6	2	1		
7	4	2	0	
8	6	4	2	0
9	8	6	3	2
10	11	8	5	3
11	14	11	7	5
12	17	14	10	7
13	21	17	13	10
14	26	21	16	13
15	30	25	20	16
16	36	30	24	19
17	41	35	28	23
18	47	40	33	28
19	54	46	38	32
20	60	52	43	37
21	68	59	49	43
22	75	66	56	49
23	83	73	62	55
24	92	81	69	61
25	101	90	77	68
26	110	98	85	76
27	120	107	93	84
28	130	117	102	92
29	141	127	111	100
30	152	137	120	109
z	1.645	1.960	2.326	2.576

För stora n är $T_\alpha \approx \dfrac{n(n+1)}{4} - z\sqrt{\dfrac{n(n+1)(2n+1)}{24}}$ där z ges i sista raden.

Hypotesen $m \geq m_0$ förkastas om $t^+ \leq T_{2\alpha}$,
Hypotesen $m \leq m_0$ förkastas om $t^- \leq T_{2\alpha}$,
Hypotesen $m = m_0$ förkastas om $t \leq T_\alpha$ (där t är det minsta av talen t^+ och t^-).

Wilcoxons rangsummetest

$U_i = W_i - \dfrac{n_i(n_i+1)}{2}$, $i = 1, 2$, där W_i är respektive rangsumma.

Kritiska värden $U_{0.10}$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2				0	0	0	1	1	1	1	2	2	3	3	3	3	4	4	4
3		0	0	1	2	2	3	4	4	5	5	6	7	7	8	9	9	10	11
4		0	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18
5	0	1	2	4	5	6	8	9	11	12	13	15	16	18	19	20	22	23	25
6	0	2	3	5	7	8	10	12	14	16	17	19	21	23	25	26	28	30	32
7	0	2	4	6	8	11	13	15	17	19	21	24	26	28	30	33	35	37	39
8	1	3	5	8	10	13	15	18	20	23	26	28	31	33	36	39	41	44	47
9	1	4	6	9	12	15	18	21	24	27	30	33	36	39	42	45	48	51	54
10	1	4	7	11	14	17	20	24	27	31	34	37	41	44	48	51	55	58	62
11	1	5	8	12	16	19	23	27	31	34	38	42	46	50	54	57	61	65	69
12	2	5	9	13	17	21	26	30	34	38	42	47	51	55	60	64	68	72	77
13	2	6	10	15	19	24	28	33	37	42	47	51	56	61	65	70	75	80	84
14	3	7	11	16	21	26	31	36	41	46	51	56	61	66	71	77	82	87	92
15	3	7	12	18	23	28	33	39	44	50	55	61	66	72	77	83	88	94	100
16	3	8	14	19	25	30	36	42	48	54	60	65	71	77	83	89	95	101	107
17	3	9	15	20	26	33	39	45	51	57	64	70	77	83	89	96	102	109	115
18	4	9	16	22	28	35	41	48	55	61	68	75	82	88	95	102	109	116	123
19	4	10	17	23	30	37	44	51	58	65	72	80	87	94	101	109	116	123	130
20	4	11	18	25	32	39	47	54	62	69	77	84	92	100	107	115	123	130	138

Kritiska värden $U_{0.05}$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2							0	0	0	0	1	1	1	1	1	2	2	2	2
3				0	1	1	2	2	3	3	4	4	5	5	6	6	7	7	8
4			0	1	2	3	4	4	5	6	7	8	9	10	11	11	12	13	14
5		0	1	2	3	5	6	7	8	9	11	12	13	14	15	17	18	19	20
6		1	2	3	5	6	8	10	11	13	14	16	17	19	21	22	24	25	27
7		1	3	5	6	8	10	12	14	16	18	20	22	24	26	28	30	32	34
8	0	2	4	6	8	10	13	15	17	19	22	24	26	29	31	34	36	38	41
9	0	2	4	7	10	12	15	17	20	23	26	28	31	34	37	39	42	45	48
10	0	3	5	8	11	14	17	20	23	26	29	33	36	39	42	45	48	52	55
11	0	3	6	9	13	16	19	23	26	30	33	37	40	44	47	51	55	58	62
12	1	4	7	11	14	18	22	26	29	33	37	41	45	49	53	57	61	65	69
13	1	4	8	12	16	20	24	28	33	37	41	45	50	54	59	63	67	72	76
14	1	5	9	13	17	22	26	31	36	40	45	50	55	59	64	69	74	78	83
15	1	5	10	14	19	24	29	34	39	44	49	54	59	64	70	75	80	85	90
16	1	6	11	15	21	26	31	37	42	47	53	59	64	70	75	81	86	92	98
17	2	6	11	17	22	28	34	39	45	51	57	63	69	75	81	87	93	99	105
18	2	7	12	18	24	30	36	42	48	55	61	67	74	80	86	93	99	106	112
19	2	7	13	19	25	32	38	45	52	58	65	72	78	85	92	99	106	113	119
20	2	8	14	20	27	34	41	48	55	62	69	76	83	90	98	105	112	119	127

Kritiska värden $U_{0.02}$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2												0	0	0	0	0	0	1	1
3						0	0	1	1	1	2	2	2	3	3	4	4	4	5
4				0	1	1	2	3	3	4	5	5	6	7	7	8	9	9	10
5			0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
6			1	2	3	4	6	7	8	9	11	12	13	15	16	18	19	20	22
7		0	1	3	4	6	7	9	11	12	14	16	17	19	21	23	24	26	28
8		0	2	4	6	7	9	11	13	15	17	20	22	24	26	28	30	32	34
9		1	3	5	7	9	11	14	16	18	21	23	26	28	31	33	36	38	40
10		1	3	6	8	11	13	16	19	22	24	27	30	33	36	38	41	44	47
11		1	4	7	9	12	15	18	22	25	28	31	34	37	41	44	47	50	53
12		2	5	8	11	14	17	21	24	28	31	35	38	42	46	49	53	56	60
13	0	2	5	9	12	16	20	23	27	31	35	39	43	47	51	55	59	63	67
14	0	2	6	10	13	17	22	26	30	34	38	43	47	51	56	60	65	69	73
15	0	3	7	11	15	19	24	28	33	37	42	47	51	56	61	66	70	75	80
16	0	3	7	12	16	21	26	31	36	41	46	51	56	61	66	71	76	82	87
17	0	4	8	13	18	23	28	33	38	44	49	55	60	66	71	77	82	88	93
18	0	4	9	14	19	24	30	36	41	47	53	59	65	70	76	82	88	94	100
19	1	4	9	15	20	26	32	38	44	50	56	63	69	75	82	88	94	101	107
20	1	5	10	16	22	28	34	40	47	53	60	67	73	80	87	93	100	107	114

Kritiska värden $U_{0.01}$

$n_1 \backslash n_2$	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
2																		0	0
3								0	0	0	1	1	1	2	2	2	2	3	3
4					0	0	1	1	2	2	3	3	4	5	5	6	6	7	8
5				0	1	1	2	3	4	5	6	7	7	8	9	10	11	12	13
6			0	1	2	3	4	5	6	7	9	10	11	12	13	15	16	17	18
7			0	1	3	4	6	7	9	10	12	13	15	16	18	19	21	22	24
8			1	2	4	6	7	9	11	13	15	17	18	20	22	24	26	28	30
9		0	1	3	5	7	9	11	13	16	18	20	22	24	27	29	31	33	36
10		0	2	4	6	9	11	13	16	18	21	24	26	29	31	34	37	39	42
11		0	2	5	7	10	13	16	18	21	24	27	30	33	36	39	42	45	48
12		1	3	6	9	12	15	18	21	24	27	31	34	37	41	44	47	51	54
13		1	3	7	10	13	17	20	24	27	31	34	38	42	45	49	53	57	60
14		1	4	7	11	15	18	22	26	30	34	38	42	46	50	54	58	63	67
15		2	5	8	12	16	20	24	29	33	37	42	46	51	55	60	64	69	73
16		2	5	9	13	18	22	27	31	36	41	45	50	55	60	65	70	74	79
17		2	6	10	15	19	24	29	34	39	44	49	54	60	65	70	75	81	86
18		2	6	11	16	21	26	31	37	42	47	53	58	64	70	75	81	87	92
19	0	3	7	12	17	22	28	33	39	45	51	57	63	69	74	81	87	93	99
20	0	3	8	13	18	24	30	36	42	48	54	60	67	73	79	86	92	99	105

För stora värden på n_1 och n_2 gäller $U_\alpha \approx \dfrac{n_1 n_2}{2} - z_\alpha \sqrt{\dfrac{n_1 n_2 (n_1 + n_2 + 1)}{12}}$.

Här är $z_{0.1} = 1.645,\ z_{0.05} = 1.960,\ z_{0.02} = 2.326,\ z_{0.01} = 2.576$.

Hypotesen $m_1 \geq m_2$ förkastas om $U_1 \leq U_{2\alpha}$,

Hypotesen $m_1 \leq m_2$ förkastas om $U_2 \leq U_{2\alpha}$,

Hypotesen $m_1 = m_2$ förkastas om $U \leq U_\alpha$ (där U är det minsta av talen U_1 och U_2).