Analytical
Methods

RSCP hin

A a a a

C : *Anal. Methods*, 2013, **5**, 5373

Analytical Methods Committee, AMCTB No 57

R c d 4 S b 2013

DOI: 10.1039/c3a 90070c

N n- ame ic t t it ical me h d , hich make fe a m ti n ab t lati n e dit ib ti n , ha e e ha been nj t l neglect ed in t he anal t ical cience . A maj ad ant age i t hat me f t hem a e im let hat t he can be ed "at t he bench."

Pa amet e n amet e ?

Anal ical scien is s generall make replica e meas remen s and rea hem as a random sample, from hich es ima es are made of he proper ies of he (h po he icall in ni e) population of meas remen s. The pop la ion mean, con dence limi s *etc.* are s all calc la ed sing he ass mp ion ha he nderl ing dis rib ion is normal (Ga ssian), i h mean μ and ariance σ^2, *i.e.* i can be s mmarised as $N(\mu,\sigma^2)$. The o erms μ and σ are he parame ers of he dis rib ion. Similarl a binomial dis rib ion is described as $B(n,\ p)$, here he parame ers n and p are respec i el he o al n mber of meas remen s and he probabili of one of he o possible o comes.

This parame er-based approach o da a handling is no essen ial, and ma no al a s be appropria e. Some imes i is kno n ha a pop la ion dis rib ion is no normal or e en close o i , so ded c ions made on he ass mp ion of normali migh be nreliable. This is par ic larl r e in cases here he same meas remen s are made on similar b non-iden ical sample ma erials of na ral origin. The an ibod le els in blood plasma samples from differen h man s bjec s are ro ghl log-normall dis rib ed, i h he addi ion of some s bjec s i h e cep ionall high le els in ario s disease s a es. Me hods ha do no make ass mp ions abo he form of he pop la ion dis rib ion are called non-parame ric or dis rib ion-free me hods. In appl ing hem he familiar approach o signi cance es ing is s ill sed. We se p a n ll h po hesis H_0 and nd he probabili of ob aining he ac al or more e reme res l s if H_0 is r e: if his probabili is er lo H_0 is rejec ed. B heir simplici makes non-parame ric me hods a rac i e e en in si a ions here more familiar es s s ch as he *t*- es migh o her ise be applied, as he e amples belo ill sho .

S me im le e am le

S ppose ha an anal ical reagen is s a ed o ha e a p ri of 99.5%, and ha s ccessi e ba ches are fo nd o ha e p ri le els of 99.2%, 99.8%, 98.9%, 99.4%, 99.1%, 99.3%, and 99.0%. Is here e idence ha he p ri of he ma erial is lo er han i sho ld be? S ch res l s are nlikel o come from a normal pop la ion (after all, he ma im m possible p ri is 100%) so a *t*- es or o her parame ric approach co ld ell be nsafe. A ke s a is ic here is he median: he n ll h po hesis is ha he da a come from a pop la ion i h a median p ri le el of 99.5%. To carr o he es e simpl s b rac his median from each of he e perimen al res l s, and no e he sign of he res l . This gi es si min s signs and one posi i e sign, *i.e.* si of he se en res l s lie belo he median. (An res l ha eq als he h po he ical median is ignored comple el). The probabili of ge ing si (or more) min s signs o of se en is pro ided b he binomial heorem, b he al es are pro ided in s a is ical ables, and can be memorised if e al a s make he same n mber of meas remen s. Here he probabili of ge ing 6 or more min s signs is 0.0625, a li le higher han he probabili le el commonl sed in signi cance es ing ($p = 0.05$), so e re ain he n ll h po hesis ha he res l s co ld come from a pop la ion i h a median p ri of 99.5%. As al a s e ha e no pro ed ha he do come from s ch a pop la ion: e ha e

[redacted]

failed o dispro e i . No e ha his is a one- ailed es , as he q es ion is he her he p ri is lo er han i sho ld be. Wi h se en meas remen s he n ll h po hesis o ld onl be rejec ed a he $p = 0.05$ le el if all se en res l s gi e min s signs hen compared i h he median al e: his o come has a probabili of onl $(1/2)^7 = 1/128$. This me hod is called he sign es , and i can be e ended o o her si a ions, s ch as comparing o se s of paired res l s, or s d ing a possible rend in a seq ence of res l s.

Ano her simple es i h man applica ions is called T ke s Q ick Tes (after John W. T ke , a major fig re in non-parame ric s a is ics and ini ial da a anal sis) or he Tail Co n Tes , he la er being a good descrip ion of i s opera ion. I is sed o compare o independen da a se s, hich need no be of he same si e. S ppose e ob ain si al es of he le el of a mospheric NO_x (μg m^{-3}) a a roadside si e: 128, 121, 117, 125, 131 and 119. A a nearb off-road si e e make si more meas remen s sing he same anal ical me hod, ob aining he res l s 120, 108, 109, 112, 114 and 110 μg m^{-3}. Is here an e idence ha he NO_x le el is lo er a he second si e han a he firs ? These o se s of res l s co ld be compared sing a (one- ailed) t- es , b he T ke approach is simpler. We simpl co n he n mber of res l s in he firs da a se ha are higher han all he al es in he second se (here are 4 of hem), and he n mber of al es in he second se ha are lo er han all hose in he firs se (5 of hem). If ei her of hese co n s is ero, he es ends a once i h he n ll h po hesis (here, ha mo ing a a from he road does no affec he NO_x le el) being accep ed. O her ise he o co n s are added oge her o pro ide he es s a is ic T (= 9 here), and his is compared i h he cri ical al e. For a one- ailed es a $p = 0.05$, T m s be grea er han or eq al o 6 if H_0 is o be rejec ed. So H_0 can be rejec ed here; he NO_x le el a he off-road si e does seem o be lo er. The meri of he T ke me hod is ha if he o al n mber of meas remen s is no more han ~20, and if he o sample si es are no grea l differen (condi ions often me in anal ical prac ice), he cri ical T al es are independen of sample si e! For he rejec ion of he n ll h po hesis in a one- ailed es he al e of T m s be ≥6, 7, 10 and 14 a $p = 0.05$, 0.025, 0.005, and 0.0005 respec i el . For a o- ailed es he corresponding cri ical al es of T are 7, 8, 11 and 15 respec i el . This remarkable fea re of he me hod means ha i can be carried o sing men al ari hme ic onl .

What n t t like?

Man non-parame ric me hods ha e been de eloped, incl ding es s analogo s o he familiar t- and F- es s, anal sis of ariance, and calibra ion and regression me hods, b despi e heir prac ical meri s onl a fe ha e fo nd fa o r in he anal ical sciences. One possible reason for his is ha mos non-parame ric me hods need a sample of a leas 6 meas remen s. Ano her reason is he gro ing pop lari of rob s me hods (AMCTB 6, 50), hich are ell s i ed o he common si a ion here he error dis rib ion is nimodal b no er differen from Ga ssian. F r hermore i is e iden ha in he o e amples abo e he f ll n merical con en of he da a is no sed. In he sign es onl he signs of he differences are co n ed, no heir magni de; and in he T ke me hod he es s a is ic is again a co n ra her han an e ac re ec ion of he n merical res l s. We migh h s e pec ha non-parame ric me hods o ld be poorer han me hods ise