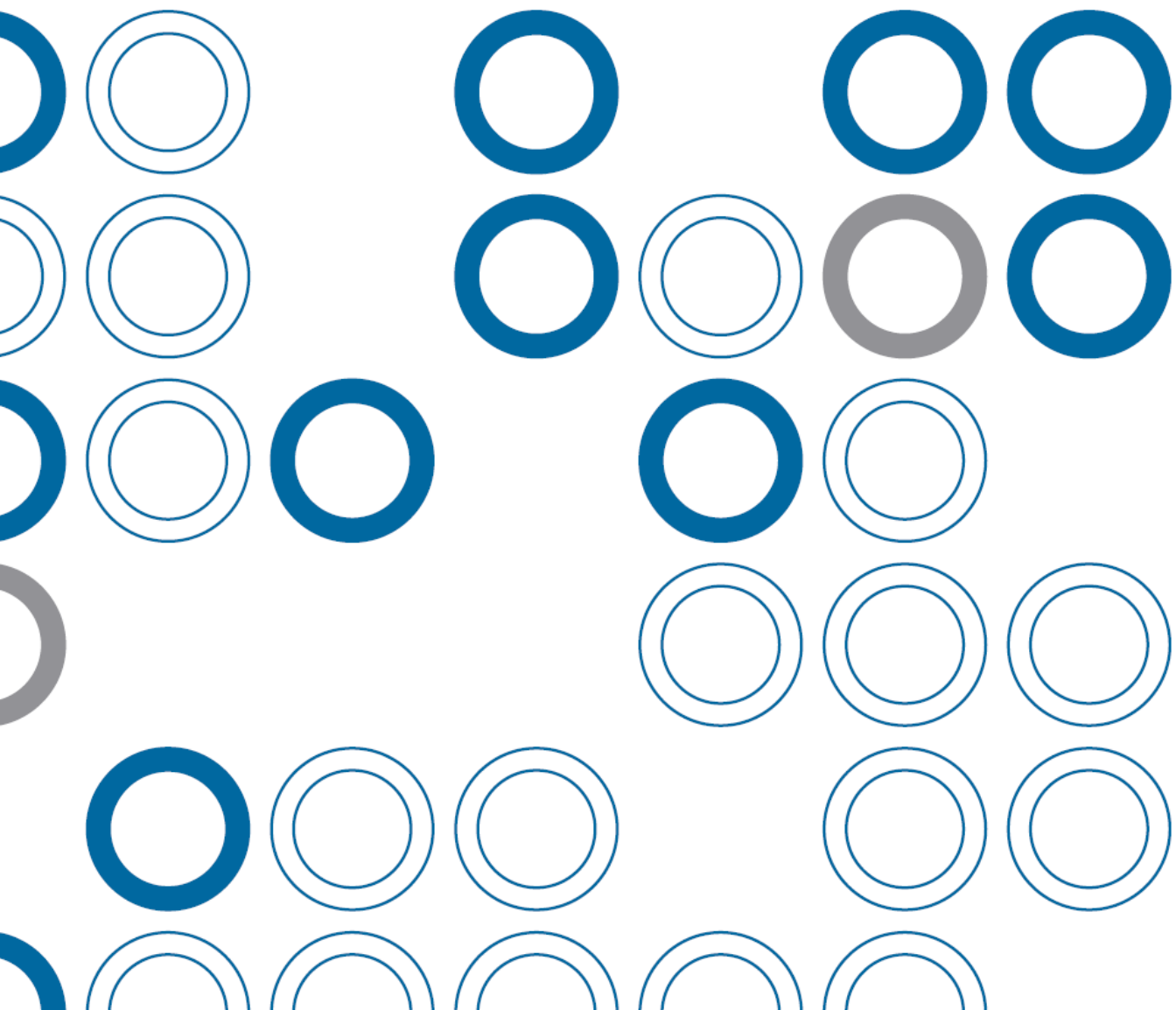


EQUALIS

Beräkningar och diagram i Equalis resultatsammanställningar

Equalis, P016, version 2.1



Innehåll

Introduktion	2
Equalis resultatsammanställningar	3
Diagram i Equalis resultatrapporter	8
Statistiska grundbegrepp	10
Referenser	13

Introduktion

Equalis arrangerar program för extern kvalitetssäkring inom hela det laboratoriemedicinska fältet. Ett flertal av programmen hanterar kvantitativa resultat, till exempel vid mätning av koncentrationen av ett visst ämne i plasma eller blod. En omgång inom ett sådant program går oftast till så att ett provmaterial delas upp i delprov som skickas ut till ett flertal deltagare som analyserar materialet och rapporterar tillbaka analysresultaten. Equalis sammanställer resultaten och skickar resultatrapporter till deltagarna med information om hur väl deras resultat överensstämmer med övriga deltagares resultat och med det förväntade svaret. Ibland specificeras ett intervall, kvalitetsmål, inom vilket resultatet bör ligga för att analyskvaliteten ska anses optimal.

Vid resultatsammanställningarna används ofta beskrivande statistik. Genom att anta att de inrapporterade analysresultaten i en omgång är ett stickprov ur en normalfördelad population (av ett oändligt antal möjliga resultat) kan analysresultaten (stickprovet) användas för att beräkna medelvärde och spridning. Ju större stickprovet är, dvs. ju fler analysresultat som ingår, desto säkrare blir skattningen av dessa statistiska mått.

I detta dokument beskrivs vilka beräkningar som utförs av Equalis i samband med resultatsammanställningarna, samt de olika typer av diagram som används i Equalis resultatrapporter för att grafiskt illustrera resultaten. Dessutom ges en kortfattad beskrivning av vissa statistiska grundbegrepp. Dokumentet riktar sig i första hand till dig som arbetar med kvalitetssäkring av laboratorieanalyser och är intresserad av att veta mer om hur Equalis gör resultatsammanställningar inom de kvalitetssäkringsprogram som hanterar kvantitativa resultat.

Equalis resultatsammanställningar

Beräkning av medelvärde och spridning

I Equalis resultatsammanställningar beräknas ofta medelvärdet och spridningen av deltagarnas resultat. Varje deltagare får ett mått på hur mycket de avviker från medelvärdet och information om hur stor spridningen är, dvs. hur samtliga deltagares (anonyma) resultat fördelat sig. Spridningen uttrycks som standardavvikelse (SD) och variationskoefficient (CV). Före beräkningen delas resultaten vanligtvis in i olika grupper beroende på vilken mätmetod deltagarna använt, så det blir möjligt att upptäcka systematiska skillnader mellan olika metoder. Vid resultatsammanställningen beräknas dels medelvärde och spridning för varje rapportgrupp separat, dels medelvärde och spridning för alla resultat tillsammans. I vissa fall kan Equalis med en kommentar i rapporten uppmana deltagarna att enbart jämföra sina resultat med resultaten från den egna rapportgruppen. Detta kan till exempel förekomma när resultaten i de separata rapportgrupperna är normalfördelade, men alla resultaten tillsammans inte är det (exempelvis på grund av metodskillnader). Om antalet deltagare i en rapportgrupp är mindre än fyra beräknas inte spridningen eftersom underlaget då anses för litet för att ge tillförlitliga resultat.

Ibland inträffar det att någon eller några deltagare rapporterar resultat som avviker på ett extremt sätt från övriga deltagares resultat, exempelvis beroende på provförväxling eller tillfälligt apparatfel. Denna typ av påtagligt avvikande resultat kallas extremvärden (outliers). För att få en säkrare beräkning av medelvärde och spridning, där inflytandet av eventuella extremvärden minimeras, använder Equalis den robusta beräkningsmetoden Algoritm A som rekommenderas i en internationell standard (ISO 13528:2022) [1]. Metoden baseras på ett antal upprepade beräkningar (iterationer) där medianvärdet används för att beräkna gränsvärden utanför vilka resultaten byts ut mot gränsvärdena. Sedan beräknas preliminärt medelvärde och SD som används för att beräkna nya gränsvärden, osv., vilket upprepas tills gränsvärdena inte längre ändras.

Algoritm A

1. Identifiera medianvärdet för samtliga originalresultat.
2. Beräkna avvikelsen från medianvärdet i absoluta tal för varje enskilt originalresultat.
3. Beräkna medianavvikelsen (= medianvärdet för avvikelserna från steg 2).
4. Beräkna preliminär SD (SD^*) enligt: $SD^* = \text{medianavvikelsen} \times 1,483$.
5. Beräkna gränsvärden $\text{medianvärdet} - 1,5 \times SD^*$ och $\text{medianvärdet} + 1,5 \times SD^*$.
6. Ersätt de resultat som ligger utanför gränserna med det närmaste gränsvärdet.
7. Beräkna preliminärt medelvärde (m^*) och preliminärt SD (SD^*) för de nya resultaten på vanligt sätt
8. Beräkna nya gränsvärden: $m^* + 1,5 \times 1,134 \times SD^*$ och $m^* - 1,5 \times 1,134 \times SD^*$.
9. Ersätt de originalresultat som ligger utanför gränserna med gränsvärdena från steg 8.
10. Beräkna nytt preliminärt medelvärde (m^*) och SD^*
11. Upprepa steg 8-10 tills det robusta medelvärdet och SD inte förändras från en iteration till nästa (i tredje signifikanta siffran). Slutligt medelvärde är medelvärdet (m^*) från sista iterationen (upprepningen) och slutlig SD är $1,134 \times SD^*$ från sista iterationen.

Jämförelse med förväntat svar

Varje deltagares resultat jämförs med ett värde som Equalis tilldelar en egenskap för ett provmaterial. Detta värde benämns "Förväntat svar". Det förväntade svaret bör vara så nära det "sanna värdet" som möjligt med de givna förutsättningarna.

Det förväntade svaret för resultat på **kvotskala** baseras på en av följande metoder:

Metod	Beskrivning
Metod 1	Ett beräknat värde baserat på en känd tillsats av komponenten till provmaterialet.
Metod 2	Ett referensvärde baserat på mätningar vid ett eller flera oberoende referenslaboratorier. Detta anses som det bästa sättet att bestämma ett åsatt värde, men är ofta kostsamt och är begränsat till de komponenter där referensmetodik finns tillgänglig. Därför används referensvärden bara ibland.
Metod 3	Ett värde från kalibrering mot referensvärdet i ett CRM (certified reference material).
Metod 4	Ett värde baserat på mätningar vid ett antal expertlaboratorier som väljs ut av Equalis.
Metod 5	Ett värde baserat på medelvärdet av rapportgruppsmedelvärden.
Metod 6	Ett rapportgruppsmedelvärde för den rapportgrupp deltagarens resultat ingår i.
Metod 7	Ett totalmedelvärde baserat på alla resultat (konsensusmedelvärde).
Metod 8	Ett värde beräknat med alternativ metod, härlett från expertkonsultation och specificerat i rapporter.

Om någon av metoderna 1-5 används kallas det för ett åsatt värde. Åsatt värde används som samlingsbegrepp för förväntat svar som inte baseras på medelvärde av deltagarnas resultat.

För mätningar som ger resultat på **ordinal/nominalskalan** kan det förväntade svaret baseras på ett av följande:

Metod	Beskrivning
Metod 1	Genom expertbedömning/analys på expertlaboratorier.
Metod 2	Genom användning av referensmaterial.
Metod 3	Från kunskap om ursprunget eller förberedelsen av provmaterialet.
Metod 4	Använda typvärdet eller medianen från deltagarresultat (medianen är endast lämplig för ordinalskalan).

Andra metoder som kan visa sig ge tillförlitliga resultat kan också användas.

Osäkerhet i förväntat svar

Osäkerhet är ett generellt begrepp som omfattar olika felkällor och variabilitet i mätningar. Det karakteriserar den samling av värden som det sanna värdet ligger inom med en viss säkerhet. Den kan t.ex. beskrivas med hjälp av en standardavvikelse/varians eller ett konfidensintervall. För resultat på ordinal/nominalskalan anges ingen osäkerhet.

Nedan redovisas hur Equalis beräknar osäkerheten i ett förväntat svar.

Metod	Beräkning av osäkerhet
Metod 1 Beräknad från känd tillsats	Beskrivs och beräknas i leverantörens dokumentation enligt generella principer för standardosäkerhet (osäkerhet i gravimetriska och volymetriska mätningar och renheten av ev. material som används i formuleringen).
Metod 2 Referenslab	Beskrivs och beräknas i leverantörens dokumentation enligt generella principer för standardosäkerhet (baserat på metodens osäkerhet samt andra osäkerheter under mätserien)
Metod 3 Värde från kalibrering mot referensvärdet i ett CRM (certified reference material)	$u_{char} = \sqrt{u_{crm}^2 + u_{\bar{d}}^2}$ <p>,där u_{crm} är angiven standardosäkerheten för CRM och $u_{\bar{d}}$ är standardosäkerheten för \bar{d} (dvs $\frac{sd}{\sqrt{p}}$).</p>
Metod 4 Medelvärde från utvalda expertlaboratorier	95% konfidensintervall för medelvärdet, $[1,96 \cdot SD / \sqrt{n}]$. SD = SD beräknat på ingående värden n = antal grupper
Metod 5 Medel av rapportgruppsmedelvärden	Samma som metod 4.
Metod 6 Totalmedelvärde	Osäkerhet beräknas: $2 * 1,25 * \frac{sd}{\sqrt{p}}$ Vilket är en utvidgad osäkerhet med täckningsfaktor $k=2$ (ca 95% konfidensintervall). sd = Robust sd från algoritim A, p = antal resultat
Metod 7 Rapportgrupps-medelvärde	Samma som metod 6
Metod 8	Ett värde beräknat med alternativ metod, härlett från expertkonsultation och specificerat i rapporter.

Bedömning av enskilda resultats kvalitet

Standardproceduren för extern kvalitetssäkring är att varje deltagares resultat jämförs med ett värde som Equalis tilldelar de komponenter som ska mätas/undersökas i provmaterialet (Förväntat svar).

Skillnaden (avvikelsen) mellan deltagarens svar och förväntat svar används sedan i bedömningen av resultatet. För bedömning av enskilda resultats kvalitet används i förväg

uppsatta kriterier i form av maximal tillåten avvikelse (acceptansgränser). Resultat inom acceptansgränsen anses ha fullgod kvalitet.

Komponenter på kvotskala

Antingen finns Equalis kvalitetsmål specificerat för komponenten och/eller så används omgångens robusta standardavvikelse som kriterium. Deltagare får själva specificera vilket kriterium som ska användas.

Equalis kvalitetsmål

Equalis kvalitetsmål finns specificerade i U040 och fastställs oftast av Equalis expertgrupp inom området. U040 finns tillgänglig på Equalis webbplats:

<https://www.equalis.se/sv/deltagare/deltagarinformation/bedomning-av-resultat/>

Det är Equalis strävan att kvalitetsmål ska införas i kvalitetssäkringsprogrammen för alla komponenter där detta är tillämpligt, och att de ska fastställas i dialog med Equalis expertgrupper och den övriga professionen.

Exempel på kvalitetsmål:

- Standard är i KM i % från förväntat svar (tex P-Albumin +/- 5%).
- Absolut gräns används för några komponenter (tex P-pH +/-0,02)
- Gräns uttryckt i SD (z-score), tex för B—Leukocyter i transfusionsmedicin (+/- 3SD)
- Sammansatt gräns för B—HbA1c (IFCC) +/- (1,5 + 1,65 x SD). Dvs en fast och en baserad på SD.

Avvikelse i SD (z-score)

Generellt kan acceptansgränsen +/-3SD användas som ersättning till +/-KM. Även acceptansgränsen +/-2SD anges i Equalis rapporter. Regler för hur dessa gränser ska användas bestämmer deltagare själva men generellt tolkas +/-2SD som en varningsgräns och +/-3SD som en åtgärdsgräns.

Dock är svagheten med z-score att gränserna beror på deltagarnas prestation, dvs ju bättre precision desto strängare blir kriteriet och tvärtom.

Komponenter på ordinal eller nominalskala

I de flesta fall finns förväntat svar för komponenten framtagen. När förväntat svar finns används kriteriet "Enligt förväntat svar", dvs om deltagarens svar är exakt samma som det förväntade svaret. Om inte det är fallet avviker svaret. I vissa program används även kriteriet "Acceptabelt svar" vilket indikerar att svaret avviker från det förväntade men att avvikelsen är av mindre art.

Utöver dessa allmänna kriterier finns även Equalis kvalitetsmål uppsatta för vissa komponenter. Exempel på kvalitetsmål:

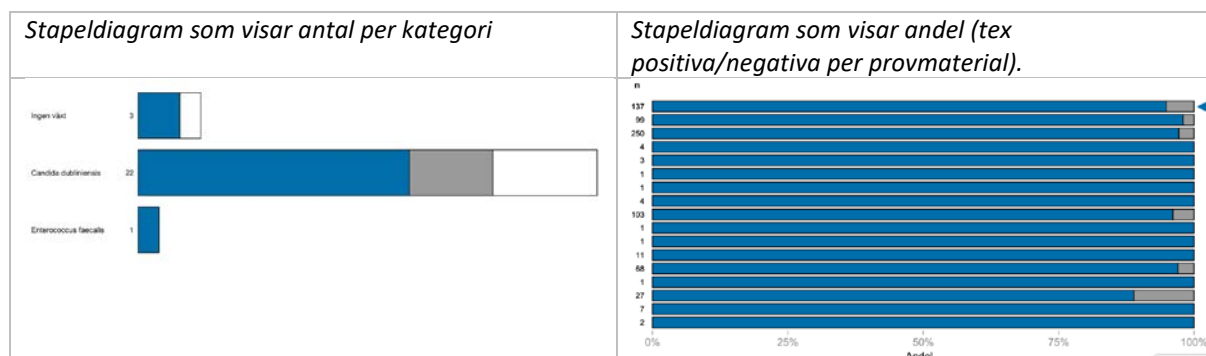
- Vanligas är "Ingen avvikelse från förväntat svar" (tex DNA-genotyp)
- Inga falskt negativa svar (tex P—HIV 1+2 (ak+ag))
- ± 1 steg från medianvärdet (tex Kvantifiering av ABO-antikroppar)

Diagram i Equalis resultatrapporter

Equalis resultatsammanställningar redovisas för deltagarna i form av olika rapporter, vars innehåll anpassas till de olika kvalitetssäkringsprogrammen. Nedan finns de vanligast använda graferna/diagrammen förklarade.

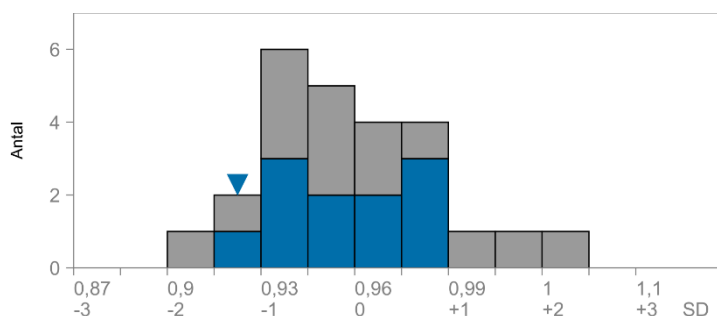
Stapeldiagram

Stapeldiagrammet visar antal per kategori, dvs. staplar där höjden är proportionell mot antalet resultat inom varje kategori. En variant används också för att visa olika svarsalternativs andel per kategori.



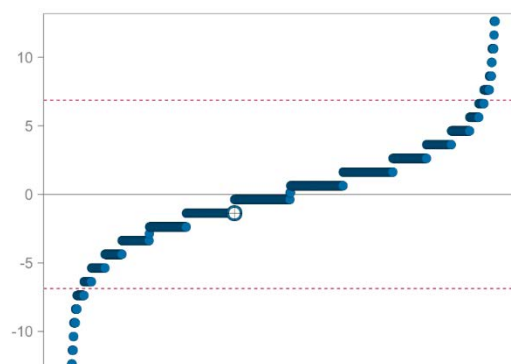
Histogram

Histogrammet är ett stapeldiagram som visar resultatens fördelning, dvs. antal resultat i olika koncentrationsintervall. Resultaten illustreras med staplar där höjden är proportionell mot antalet resultat inom varje intervall.



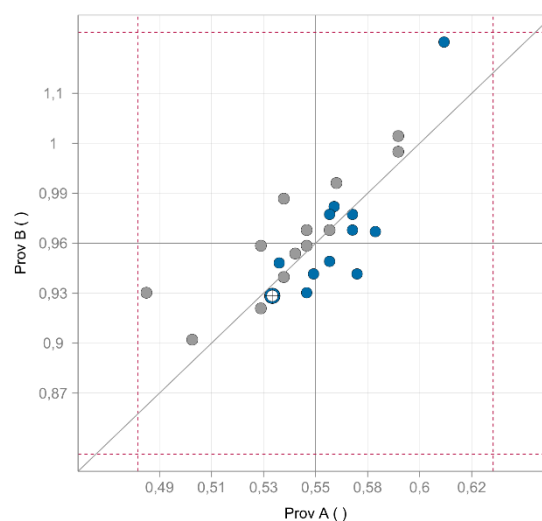
Avvikelsediagram

Denna diagramtyp illustrerar hur mycket deltagarnas resultat avviker från ett givet värde. Samtliga deltagares avvikelser visas i stigande ordning längs x-axeln. På y-axeln visas avvikelsens storlek. I diagrammet kan kvalitetsmål ritas in som linjer vid i förväg bestämda acceptansgränser ($\pm 5\%$ i exemplet till höger) inom vilka analyskvaliteten anses vara optimal.



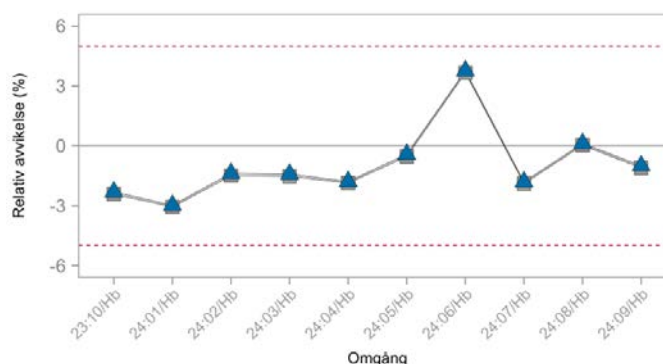
Youdendiagram

Youdendiagrammet kan användas för att visa analysresultat för två olika prover i samma diagram. Analysresultaten från de två provmaterialen A och B (med olika koncentration) avsätts mot varandra, med analysresultaten för prov A på den ena axeln och för prov B på den andra axeln. Varje deltagande laboratorium representeras i diagrammet av en punkt vars koordinater utgörs av analysresultaten för prov A och B. Tillsammans bildar alla deltagarnas resultat en punktsvärm, vars utseende kan vara till hjälp för att upptäcka eventuella samband mellan avvikande analysresultat och koncentrationsnivå. Om kvalitetsmål finns tillgängliga ritas dessa ut för båda proven som röda linjer.



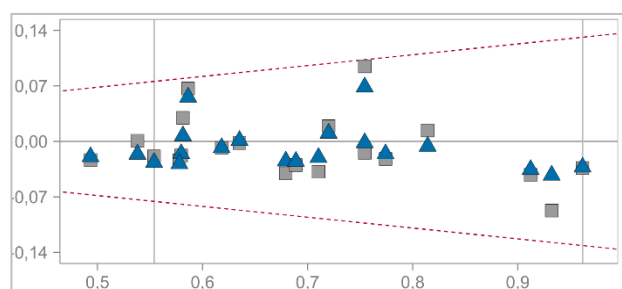
Shewhartdiagram

Ett Shewhartdiagram används för att följa en enskild deltagares analyskvalitet över tid. Diagrammet har en x-axel för tid (t.ex. datum eller omgångsnummer för ett kontrollutskick) och en y-axel för avvikelsen mellan enskilda analysresultat och ett givet värde. Även i Shewhartdiagrammet kan kvalitetsmål ritas in i form av i förväg bestämda acceptansgränser.



Nivådiagram

Visar hur deltagarens resultat har avvikit från ett givet värde i förhållande till koncentration av komponenten (nivån) i kontrollmaterialet. Avvikelser från de senaste analysomgångarna visas i diagrammet och aktuell omgång markeras med en grå linje. Avvikelser som varierar vid olika nivåer (koncentrationer) kan vara tecken på att avvikelserna är koncentrationsberoende.



Statistiska grundbegrepp

Fördelning

Fördelningen av ett antal värden beskriver hur många gånger varje värde förekommer. Vid sammanställning av analysresultat från ett kontrollutskick beskrivs ofta analysresultatens fördelning med ett histogram.

Medelvärde

Medelvärdet används för att ange fördelningens läge (koncentrationsnivå). Det beräknas genom att man summerar alla enskilda observationsvärden och dividerar summan med antalet värden.

Matematiskt uttrycks detta som:
$$m = \frac{\sum x}{n}$$

där m är medelvärdet, n är antalet värden och $\sum x$ är summan av alla (n) värden.

Medianvärde

Medianvärdet är ett annat sätt att ange fördelningens läge (koncentrationsnivå). Det är ett centralmått, kring vilket övriga värden ligger fördelade med lika många värden på vardera sidan. Medianvärdet är särskilt väl lämpat för att ange läget av sneda fördelningar. För helt symmetriska fördelningar överensstämmer medianvärdet med medelvärdet.

Standardavvikelse

Standardavvikelsen (SD) ger upplysning om resultatens spridning, dvs. fördelningens bredd. Den är ett mått på hur mycket resultaten avviker från medelvärdet.

Matematiskt uttrycks detta som:
$$SD = \sqrt{\frac{\sum (x - m)^2}{n - 1}}$$

där SD är standardavvikelsen, $\sum (x - m)^2$ är summan av alla kvadrerade avvikelser från medelvärdet (m) och n är antalet avvikelser.

Variationskoefficient

Variationskoefficienten (CV) är ett normaliserat sätt att beskriva resultatens spridning, uttryckt som standardavvikelsen (SD) i förhållande till medelvärdet (m). Variationskoefficienten uttrycks oftast i procent.

Används för att jämföra spridning på olika nivåer/koncentrationer (SD varierar beroende på medelvärdets storlek).

Matematiskt uttrycks detta som:
$$CV = \frac{SD}{m} \cdot 100$$

Normalfördelning

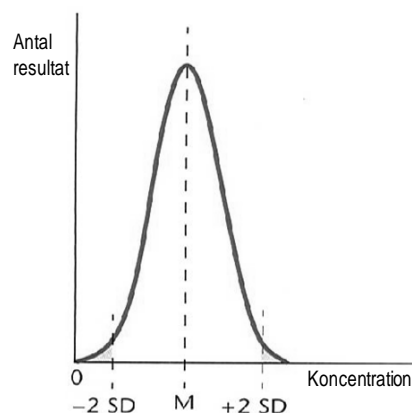
Normalfördelningen är en statistisk fördelning som är symmetrisk kring sitt medelvärde. Fördelningen karakteriseras av att cirka 68 % av värdena ligger inom ± 1 SD, cirka 95 % inom ± 2 SD och cirka 99,7 % inom ± 3 SD från medelvärdet.

Fördelningen uppvisar en klockliknande form (Gausskurva) som kan beräknas matematiskt med medelvärdet (M) och standardavvikelsen (SD).

Fördelningen är den vanligaste approximationen av verkligheten och en förutsättning för Equalis statistik för resultat på kvotskalan.

Exempelvis kan laboratoriernas resultat i ett externt kvalitetsäkringsprogram betraktas som ett stickprov från en population med oändligt antal möjliga resultat. Förutsatt att antalet resultat i stickprovet är tillräckligt stort kan stickprovets fördelning användas för att bedöma hur populationens fördelning ser ut.

Om stickprovsfördelningen är symmetrisk omkring medelvärdet gäller detta sannolikt också för populationsfördelningen. Populationsfördelningen antar då en klocklik form enligt diagrammet ovan som visar antal resultat vid varje koncentration.



Konfidensintervall (intervallskattning)

Intervallskattning baserar sig på en procentsats som definierar hur stor sannolikheten är att resultatet befinner sig inom ett intervall från en punktskattning (tex medelvärde). Denna procentsats kallas konfidensgrad. Om vi har en konfidensgrad på 95% kommer 5% av fallen att hamna utanför intervallet.

Referenser

1. ISO 13528:2022 Statistical methods for use in proficiency testing by interlaboratory comparison, Annex C.