MISSI dokumentasjon

Oversikt over modeller og beregninger for Metodeguidens Integrerte Software for Statistikk og Innsikt (**MISSI**).

Sist oppdatert 28.04.2025

Innhold

Deskriptiv statistikk

Deskriptiv statistikk gir en sammenfattet oversikt over fordelingen i et datasett. Den brukes til å beskrive sentraltendens, variasjon og fordelingsegenskaper i én eller flere variabler før man går videre til inferensielle analyser.

Verktøyet beregner antall gyldige verdier, gjennomsnitt, standardavvik, minimum og maksimum, range, 95% konfidensintervall (CI), interkvartilområde (IQR) og varians.

Formler og metode

Gjennomsnitt (mean)

Dette er et mål for sentraltendens, og er det aritmetiske snittet av alle gyldige (ikke-manglende) verdier.

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Standardavvik (standard deviation, SD)

Verktøyet bruker ***n - 1*** i nevneren, altså korrigert for utvalg (sample SD). Denne estimerer populasjonens standardavvik basert på et utvalg, i tråd med Bessel's korreksjon. SPSS bruker samme formel.

$$s = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}$$

Median, minimum og maksimum

Identifiserer den laveste og høyeste verdien i datasettet. Medianen er den midterste verdien når data er sortert. Hvis ***n*** er partall, tas snittet av de to midterste verdiene.

95% konfidensintervall (CI)

$$\bar{x} \pm t_{\alpha/2,\,df} \cdot \frac{s}{\sqrt{n}}$$

Angir et område som med 95% sannsynlighet inneholder det sanne gjennomsnittet i populasjonen. Det beregnes ved å ta gjennomsnittet ± t-verdi ganget med standardfeilen. Bredden på intervallet reflekterer usikkerheten

s: Standardavvik

n: antall gyldige observasjoner

$t_{\alpha/2,\,df}$: Kritisk t-verdi for valgt konfidensnivå og frihetsgrader

$\bar{x}$: Gjennomsnittet av observasjonene

Interkvartilområde (IQR)

Dette intervallet dekker de midterste 50 % av verdiene i datasettet og er robust mot ekstreme verdier. Beregnes som **Q3** minus **Q1**.

Varians

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Varians er kvadratet av standardavviket og representerer den gjennomsnittlige kvadrerte avstanden fra gjennomsnittet.

Design og avvik fra andre statistikkprogram og designvalg

- **Manglende verdier** blir filtrert ut før alle beregninger, lik SPSS.
- **Standardavvik** er estimert med samme formler som SPSS (inkl. Bessel's korreksjon og justert momentmetode).
- **Median** og kvartiler beregnes ved interpolasjon dersom det trengs, men for enkelhet er medianen alltid entydig vist.
- **IQR:** Verktøyet beregner kvartiler i tråd med interpolasjonsmetoden til SPSS, som følger en modifisert versjon av Tukey's hinges.
- **Varians:** Verktøyet benytter *Bessel's korreksjon* (***n − 1*** i nevneren), som er standard for estimering av populasjonsvarians fra et utvalg. Dette er identisk med SPSS sin beregning av varians.
- **Ingen forskjeller i metodevalg** for denne modulen – verdiene vil i praksis matche SPSS ($\pm$ numerisk avrundingsfeil ved svært små datasett).

Sammenheng (korrelasjonsanalyse)

Korrelasjonsanalyse brukes for å undersøke lineære sammenhenger mellom to kontinuerlige variabler. Verktøyet beregner følgende for hver variabelkombinasjon:

- Pearson's korrelasjonskoeffisient (r)
- Spearman's ρ (rho) – måler rangbasert (monoton) sammenheng
- p-verdi for signifikans
- Regresjonsligning

---

1. Pearson's r og Spearman's ρ (rho)

Verktøyet bruker den klassiske formelen for sample Pearson-korrelasjon:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2} \cdot \sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}}$$

Dette kan også uttrykkes som:

$$t = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{med df} = n - 2$$

Beregning av Spearman's ρ

Verktøyet bruker denne formelen når det ikke er tied ranks, hvor d_i er forskjellen i rangene mellom x_i og y_i.

$$\rho = 1 - \frac{6\sum d_i^2}{n(n^2-1)}$$

Ved tied ranks brukes standardformelen for Pearson's r, men på rangerte data. Rangeringen skjer automatisk ved «jStat.rank()».

p-verdi

Test av nullhypotesen (ingen korrelasjon i populasjonen) gjennomføres med følgende test-statistikk:

$$t = \frac{r \cdot \sqrt{n-2}}{\sqrt{1-r^2}} \quad \text{med df} = n-2$$

p-verdien beregnes som:

$$p = 2 \cdot P(t_{n-2} > |t|)$$

Parret t-test (avhengige målinger)

Parret t-test brukes når to målinger er gjort på samme enheter (f.eks. pre–post), og vi ønsker å teste om det er en signifikant endring. Verktøyet rapporterer gjennomsnitt (pre, post og differanse, standardavvik, p-verdi og effektstørrelse (Cohen's d og Hedges' g) via følgende beregninger:

Gjennomsnitt og SD

$$\bar{D} = \frac{1}{n}\sum_{i=1}^{n} D_i \quad , \quad s_D = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(D_i - \bar{D})^2}$$

Standardfeil og t-verdi

$$SE = \frac{s_D}{\sqrt{n}} \quad , \quad t = \frac{\bar{D}}{SE}$$

Frihetsgrader og p-verdi

p-verdien beregnes med Student's t-fordeling via jstat.studentt.cdf()

$$df = n - 1 \quad , \quad p = 2 \cdot P(t_{df} > |t|)$$

Cohen's d (effektstørrelse)

Effektstørrelsen beregnes via den klassiske definisjonen for parret t-test, også brukt i SPSS.

$$d = \frac{\bar{D}}{s_D}$$

Hedges' g

Hedges' g korrigerer Cohen's d for små utvalg:

$$g = d \cdot \left(1 - \frac{3}{4n - 5}\right)$$

Pooled SD og SD av differanser

Verktøyet tillater valg av alternativ beregning der man bruker **pooled SD** fra pre og post, i stedet for **SD av differansene**. SD av differanser tar hensyn til korrelasjonen mellom målingene, noe som gir lavere standardfeil og høyere teststyrke. Pooled SD antar uavhengige målinger.

$$s_{pooled} = \sqrt{\frac{(s_X^2 + s_Y^2)}{2}} \quad \text{og} \quad d_{pooled} = \frac{\bar{D}}{s_{pooled}}$$

Repeterte målinger ANOVA (én faktor)

Repeterte målinger ANOVA brukes når du har én gruppe målt tre eller flere ganger på samme variabel. Det er en utvidelse av parret t-test til flere tidspunkter (eller betingelser).

Modell og logikk

Verktøyet estimerer en **within-subjects variansanalyse** basert på klassisk ANOVA-partisjonering: «Total Varians=Mellom-tidspunkter+Person+Feil».

Modellen ser slik ut:

$$Y_{ij} = \mu + \tau_j + \pi_i + \varepsilon_{ij}$$

Hvor:

Y_{ij}: score for person i på tid j

μ: grand mean

τj: effekt av tid j (behandlingsbetingelse)

π_i: effekt av person i

ε_{ij}: residual (feil)

Beregning i verktøyet

Total Sum of Squares (SST)

$$SS_{Total} = \sum_{i=1}^{N}\sum_{j=1}^{K}(Y_{ij} - \bar{Y})^2$$

K er antall tidspunkter, N er antall personer, og $\bar{Y}$ er overall mean.

Mellom-tidspunkter (Treatment)

$$SS_{Time} = \sum_{j=1}^{K} n(\bar{Y}_{.j} - \bar{Y})^2$$

$\bar{Y}$: gjennomsnitt på tid j

n: antall personer

Person-effekt

$$MS_{Time} = \frac{SS_{Time}}{df_{Time}}, \quad df_{Time} = K - 1$$

Yi: personens gjennomsnitt over tid. Brukes til å forklare "blocking"-effekten som fjernes i repeated design.

$$SS_{Subjects} = \sum_{i=1}^{N} K(\bar{Y}_{i.} - \bar{Y})^2$$

Feil (residual)

$$SS_{Error} = SS_{Total} - SS_{Time} - SS_{Subjects}$$

$$MS_{Error} = \frac{SS_{Error}}{df_{Error}}, \quad df_{Error} = (N-1)(K-1)$$

F-verdi og p

Tolker hovedvirkningen av tid.

$$F = \frac{MS_{Time}}{MS_{Error}}$$

$$p = P(F_{df_{Time}, df_{Error}} > F)$$

Effektstørrelse

Partial Eta^2:

$$\eta^2 = \frac{SS_{Time}}{SS_{Time} + SS_{Error}}$$

Bonferroni Post-hoc

Hvis F-testen er signifikant, tester verktøyet alle parvise forskjeller. P-verdien beregnes ved t-test og justeres med Bonferroni-korreksjon: $\mathbf{p}_{adj} = p \cdot k$ Verktøyet bruker **klassisk RM-ANOVA** uten korreksjon for sferisitet (siden Mauchly's test ikke er tilgjengelig uten tilgang til matriser).

Dette er akseptabelt når tidspunktene er få (3–5) og det ikke er stor variansforskjell mellom tidspunktene.

Uavhengig t-test

Formål

En **uavhengig t-test** (også kalt *independent samples t-test*) undersøker om to **uavhengige grupper** har signifikant forskjellig gjennomsnitt på en kontinuerlig variabel.

Statistisk modell

Modellen antas som:

$$Y_{ij} = \mu_j + \varepsilon_{ij}$$

Hvor $\in \{1,2\}$ er gruppe 1 og 2, μ_j : middelverdi i gruppe j og $\varepsilon_{ij} \sim N(0, \sigma^2)$ er residual.

Hypotesetest for nullhypotese: $\mu_1 = \mu_2$

Alternativ: $\mu_1 \neq \mu_2$

Beregninger i verktøyet

Verktøyet bruker to versjoner, avhengig av om variansen antas lik:

Lik varians (standard pooled t-test)

$$t = \frac{\bar{X}_1 - \bar{X}_2}{s_p \cdot \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}$$

$$s_p^2 = \frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n_1 + n_2 - 2}$$

$\bar{X_j}$: gjennomsnitt i gruppe j

s^2_j : varians i gruppe j

n_j : antall i gruppe j

df $= n_1 + n_2 - 2$

Ulik varians (Welch's t-test) - Dersom Levene-testen (eller visuell vurdering) tyder på ulik varians:

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

$$df = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{(s_1^2/n_1)^2}{n_1 - 1} + \frac{(s_2^2/n_2)^2}{n_2 - 1}} \quad \text{(Welch-Satterthwaite)}$$

Verktøyet bruker **Welch's t-test som standard**, fordi den er mer robust ved ulik gruppestørrelse og ulik varians – akkurat som SPSS når "Equal variances not assumed" er valgt.

Effektstørrelse: Cohen's d

$$d = \frac{\bar{X}_1 - \bar{X}_2}{s_p} \quad \text{(ved lik varians)}$$

$$d = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{\frac{s_1^2 + s_2^2}{2}}} \quad \text{(ved ulik varians – verktøyets standard)}$$

Effektstørrelse: Hedges' g

Hedges' g er en justert versjon av Cohen's d som korrigerer for skjevhet i små utvalg. Den anbefales særlig når $n < 20$. Formelen er $g = J * d$ hvor g er Hedges' g, d er Cohen's d og J kan uttrykkes slik:

$$J = 1 - \frac{3}{4df - 1}$$

J er en *bias correction factor* som drar Cohen's d litt ned for små utvalg. For høye frihetsgrader vil J være nær 1, og dermed d og g være veldig like.

Enveis ANOVA (One-Way Analysis of Variance)

En enveis ANOVA benyttes for å teste om tre eller flere grupper har signifikant forskjellige gjennomsnitt på én kontinuerlig avhengig variabel. Enveis ANOVA sammenligner variasjonen *mellom gruppene* med variasjonen *innenfor gruppene* for å avgjøre om gruppetilhørighet har en signifikant effekt.

Verktøyet gjør følgende:

1. Grupperer data etter valgt gruppekolonne
2. Beregner alle nødvendige SS, MS, df, F og p
3. Viser oversiktstabell med:
    - F-verdier og p-verdier for gruppeeffekt
    - Eta^2 som mål på effektstørrelse
4. Utfører Bonferroni-justert post-hoc ved signifikant gruppeeffekt

---

Modell og formler

Den lineære modellen er:

$$Y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

hvor:

- Y_{ij}: Scoren til person j i gruppe i
- μ: Totalt gjennomsnitt
- τ_i: Effekt av gruppe i (avvik fra totalt gjennomsnitt)
- ϵ_{ij}: Tilfeldig feil (normalfordelt, uavhengig, varians σ^2)

Beregninger

Sum of Squares (SS).

Total varians:

$$SS_{\text{total}} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2$$

Mellom grupper:

$$SS_{\text{between}} = \sum_{i=1}^{k} n_i (\bar{Y}_i - \bar{Y})^2$$

Innen grupper (feil):

$$SS_{\text{within}} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y}_i)^2$$

$$SS_{\text{total}} = SS_{\text{between}} + SS_{\text{within}}$$

Frihetsgrader:

$$df_{\text{between}} = k - 1$$

$$df_{\text{within}} = N - k$$

Mean squares:

$$MS_{\text{between}} = \frac{SS_{\text{between}}}{df_{\text{between}}}$$

$$MS_{\text{within}} = \frac{SS_{\text{within}}}{df_{\text{within}}}$$

F-verdi:

$$F = \frac{MS_{\text{between}}}{MS_{\text{within}}}$$

Effektstørrelse:

Verktøyet beregner **eta²** (η^2), som måler andelen av total varians som forklares av gruppetilhørighet:

$$\eta^2 = \frac{SS_{\text{between}}}{SS_{\text{total}}}$$

Post-hoc (Bonferroni)

Ved signifikant ANOVA-test, brukes Bonferroni-justerte t-tester for å sammenligne gruppene parvis. Effektstørrelser for disse kan vises som **Cohen's d** eller **Hedges' g**:

Cohen's d:

$$d = \frac{\bar{X}_1 - \bar{X}_2}{SD_{\text{pooled}}}$$

Hedges' g:

$$g = d \cdot \left(1 - \frac{3}{4(n_1 + n_2) - 9}\right)$$

- Verktøyet benytter vanlig OLS-modellering og standardformler for enveis ANOVA.
- Det er testet mot SPSS og gir tilsvarende resultater så lenge antakelser er møtt (normalfordeling, lik varians).
- Ved store forskjeller i gruppestørrelser eller variansbrudd kan resultatene avvike. I slike tilfeller anbefales Welch ANOVA (ikke implementert per nå).

ANCOVA – Analyse av kovarians

ANCOVA kombinerer regresjon og ANOVA for å sammenligne gruppeforskjeller samtidig som man kontrollerer for et kontinuerlig kovariat. Det er en metode for å *justere* for variasjon forklart av en tredje variabel før man vurderer gruppeforskjeller.

Modell

Den lineære modellen er:

$$Y_{ij} = \mu + \tau_i + \beta(X_{ij} - \bar{X}) + \epsilon_{ij}$$

hvor:

- Y_{ij}: Avhengig variabel for person j i gruppe i
- μ: Generell intercept
- τ_i: Effekt av gruppe i
- X_{ij}: Kovariatverdi
- β: Effekt av kovariat
- ϵ_{ij}: Feilledd (normalfordelt)

Kovariatverdiene blir sentralisert (mean-centered) i mange implementasjoner (også i SPSS), men dette er ikke et krav så lenge designet er balansert.

Beregning og formler

Verktøyet estimerer og sammenligner fire modeller (OLS-baserte):

Modell	Innhold
M1	Kun kovariat
M2	Kun gruppe
M3	Gruppe + kovariat
M4	Gruppe + kovariat + interaksjon (gruppe × kovariat)

SS (sum of squares):

For alle modeller beregnes via OLS.

$$SS_{\text{error}} = \sum (Y_i - \hat{Y}_i)^2$$

Deretter beregnes:

- Gruppeeffekt:

$$SS_{\text{group}} = SS_{\text{error,M1}} - SS_{\text{error,M3}}$$

- Kovariat:

$$SS_{\text{covariate}} = SS_{\text{error,M2}} - SS_{\text{error,M3}}$$

- Interaksjon:

$$SS_{\text{interaction}} = SS_{\text{error,M3}} - SS_{\text{error,M4}}$$

Frihetsgrader:

$df_{\text{group}} = \boldsymbol{k}$ - 1

$df_{\text{kovariat}} = 1$

$df_{\text{interaksjon}} = (\boldsymbol{k} - 1)$

$df_{\text{error}} = \boldsymbol{N} - \boldsymbol{p}$ hvor p er antall parameter i modellen

F-verdi og effektstørrelse:

Partial eta²:

$$F = \frac{MS_{\text{effect}}}{MS_{\text{error}}} \quad \text{der } MS = \frac{SS}{df}$$

$$\eta^2 = \frac{SS_{\text{effect}}}{SS_{\text{total}}}$$

R²:

$$R^2_{\text{adj}} = 1 - \left(\frac{SS_{\text{error}}/df_{\text{error}}}{SS_{\text{total}}/(N-1)} \right)$$

Interaksjonstesting

Verktøyet tester eksplisitt **gruppe × kovariat**-interaksjon. Hvis denne er signifikant:

- **ANCOVA er ikke gyldig**, da forutsetningen om lik kovariat-effekt i alle grupper brytes.
- I slike tilfeller anbefales separate regresjonsanalyser per gruppe.

EM Means

Verktøyet beregner **Estimated Marginal Means (EM Means)** for hver gruppe, justert for gjennomsnittlig kovariatverdi:

$$\hat{Y}_g = \hat{\mu} + \tau_g + \beta(\bar{X}_{\text{cov}} - \bar{X})$$

95 % konfidensintervall estimeres basert på modellens feilledd og kovariansmatrise $(\boldsymbol{X^TX})^{-1}$

Post-hoc

Verktøyet kjører Bonferroni-korrigerte parvise sammenligninger mellom grupper basert på EM Means. Standardfeil beregnes basert på modellens MSE og varianser for hver gruppe. Som i SPSS blir p-verdien multiplisert med antall observasjoner heller enn å dele alpha på antall observasjoner.

Validering:

- Verktøyet gir resultater lignende SPSS (ANCOVA med Fixed Factors og covariate).
- SPSS tester interaksjon hvis man aktiverer *"custom model"* og legger inn gruppe × kovariat. Dette verktøyet viser interaksjon automatisk.
- Både SPSS og verktøyet bruker OLS-regresjon i bunn, men SPSS sentraliserer ofte kovariat for å redusere multikollinearitet – noe som ikke påvirker F-verdier eller tolkning.
- Verktøyet gir full kontroll over modellstrukturen og eksplisitt rapportering av alle delmodeller.

Toveis ANOVA

Toveis ANOVA (2x2 eller 2x3, osv.) benyttes når man ønsker å undersøke hvordan to uavhengige kategoriske variabler (faktorer) påvirker en kontinuerlig utfallsvariabel, samt om det finnes en interaksjon mellom de to faktorene. Dette kan for eksempel være kjønn (mann/kvinne) × Betingelse (eksperiment/kontroll), eller gruppe (A, B, C) × Tidspunkt (morgen/kveld).

Modellen beregner hovedvirkning av faktor A, hovedvirkning av faktor B og interaksjon mellom A × B.

Total varians splittes i fire komponenter:

$$SS_{Total} = SS_A + SS_B + SS_{AB} + SS_{Error}$$

Frihetsgrader (df):

$$df_A = a - 1, \quad df_B = b - 1, \quad df_{AB} = (a-1)(b-1), \quad df_{Error} = N - ab$$

Mean Squares og F-verdier:

$$MS_A = \frac{SS_A}{df_A}, \quad F_A = \frac{MS_A}{MS_{Error}}$$

$$MS_B = \frac{SS_B}{df_B}, \quad F_B = \frac{MS_B}{MS_{Error}}$$

$$MS_{AB} = \frac{SS_{AB}}{df_{AB}}, \quad F_{AB} = \frac{MS_{AB}}{MS_{Error}}$$

Verktøyet rapporterer **Partial Eta²** for hver effekt:

$$\eta^2_{partial} = \frac{SS_{effekt}}{SS_{effekt} + SS_{Error}}$$

R^2 og justert R^2 - Tolkning som forklaringsgrad for hele modellen:

$$R^2 = 1 - \frac{SS_{Error}}{SS_{Total}} \quad , \quad \text{Adjusted } R^2 = 1 - \left(\frac{1 - R^2}{df_{Total}/df_{Error}} \right)$$

Metodevalg og forskjeller fra SPSS

- Verktøyet bruker lineær regresjon og dummy-koding for estimering – tilsvarer «Type III sums of squares» i SPSS.
- Resultatene fra verktøyet vil være lignende med SPSS gitt lik datarensing og faktornivåkoding.
- Post-hoc-tester er basert på Estimated Marginal Means (EM Means) og Bonferroni-korreksjon.
- EM Means og interaksjonsplott gir ekstra innsikt ved signifikant interaksjon.

Verktøyet benytter en modellbasert regresjonstilnærming til å estimere sum of squares og effekter i toveis ANOVA. Dette er i tråd med moderne analytiske standarder og er i samsvar med lineær modell-teori. Tilnærmingen har noen forskjeller fra det som SPSS typisk rapporterer, hovedsakelig knyttet til hvordan sum of squares (SS) beregnes.

Type I, II og III sum of squares

- SPSS bruker som standard Type III SS når faktorer er modellert som faste effekter.
- Verktøyet bruker regresjonsbasert SS, tilsvarende Type I eller Type II, avhengig av faktorrekkefølge og modellstruktur.

Forskjellen består i hvordan variansen tilskrives faktorene når det er avhengighet mellom dem (f.eks. ved ulik gruppestørrelse eller ubalansert design).

Fordel med verktøyets metode:

- Mer transparent og matematisk konsistent med regresjonsanalyse.
- Identisk med output fra R når man bruker lm() uten eksplisitt Type III spesifikasjon.
- Krever ikke ekstra korreksjon for ubalanse hvis design er fullstendig og balansert.

Numeriske forskjeller

I praksis kan forskjeller oppstå i:

- **SS og F-verdier for hovedvirkninger** ved ubalansert design
- **Interaksjonsleddet** (A × B), hvis Type III brukes i SPSS

Men disse forskjellene er numerisk små og handler primært om fordelingsreglene for variasjon – ikke selve modellen. Begge tilnærminger tester de samme hypotesene, og ingen av dem er «feil», men må forstås i konteksten av modellvalg.

Mixed ANOVA

Mixed ANOVA (også kalt split-plot ANOVA eller mixed design ANOVA) benyttes når man har en mellomgruppefaktor (f.eks. ulike grupper) og en repetert måling innen gruppe (f.eks. pre og post).

Modelloppsett

En typisk Mixed ANOVA-modell inkluderer:

- **Hovedvirkning av gruppe** (mellomgruppesammenligning)
- **Hovedvirkning av tid** (repetert måling)
- **Interaksjon mellom tid og gruppe** (varierer utviklingen mellom gruppene?)

Modell og beregninger

For hver deltaker i i gruppe g ved tidspunkt T:

$$Y_{igt} = \mu + \alpha_g + \beta_t + (\alpha\beta)_{gt} + \epsilon_{igt}$$

Hvor:

- Y_{igt}: Verdi for deltaker i, gruppe g, tid t
- μ: Total gjennomsnitt
- α_g: Effekt av gruppe
- β_t: Effekt av tid
- $(\alpha\beta)_{gt}$: Interaksjonseffekt mellom gruppe og tid
- ϵ_{igt}: Feilledd

For hver effekt beregnes F-verdi slik:

$$F = \frac{MS_{effekt}}{MS_{feil}}$$

MS = SS / df

SS: Sum of squares

df: Frihetsgrader

Partial Eta² (Effektstørrelse) gir andelen forklart varians for en gitt effekt, uavhengig av andre effekter.

$$\eta_p^2 = \frac{SS_{effekt}}{SS_{effekt} + SS_{feil}}$$

Verktøyet beregner:

- F-verdier og p-verdier for gruppe, tid og interaksjon
- Partial eta² for effektstørrelse
- Post-hoc sammenligninger av grupper (via EM Means)
- Justerte gjennomsnitt (EM Means) for kombinasjoner av tid og gruppe
- Bonferroni-korrigerte p-verdier i post-hoc-analyser

Avvik fra SPSS og begrunnelse

Verktøyet bruker en lineær modell-tilnærming (basert på jStat og math.js) i stedet for SPSS sin GLM (General Linear Model). Det kan føre til små forskjeller i p-verdier og F-verdier, spesielt ved:

- Ulike antall observasjoner i grupper
- Mismatch mellom within-subject design og antagelser i modellen

Begrunnelse for metodevalg:

- Full åpenhet om beregningsgrunnlag
- Kan brukes i nettleser uten SPSS
- Transparent og repeterbar statistikk
- Gode approksimasjoner til klassisk mixed ANOVA

Wilcoxon Signed-Rank Test

Tester om medianen for forskjellen mellom to sammenkoblede målinger er forskjellig fra null. Dette er en ikke-parametrisk analog til parret t-test.

Testen brukes ved to målinger på samme enhet (f.eks. pre vs. post), når data er ikke normalfordelt (verken forskjeller eller råscore).

Beregningsmetode

1. Beregn differanser ($D_i = Y_i - X_i$) hvor X_i og Y_i er de to målingene for deltaker i.
2. Ekskluder nullforskjeller ved å fjerne alle ($Di = 0$)
3. Beregner absoluttverdi og ranger D_i fra lavest til høyest og deler like verdier (ties) likt.
4. Tildeler fortegn og gir hvert rangert tall samme fortegn som det opprinnelige D_i
5. Summerer positive og negative rangsummer:

$$W^+ = \sum \text{positive ranks}, \quad W^- = \sum \text{negative ranks}$$

6. Teststatistikk:

$$W = \min(W^+, W^-)$$

7. Beregner p-verdi ved å bruke eksakt p-verdi for små utvalg ($n \leq 20$), ellers benyttes z-approksimasjon:

$$z = \frac{W - \mu_W}{\sigma_W}, \quad \mu_W = \frac{n(n+1)}{4}, \quad \sigma_W = \sqrt{\frac{n(n+1)(2n+1)}{24}}$$

Effektstørrelse (r)

$$r = \frac{z}{\sqrt{n}}$$

Hva rapporterer verktøyet?

- Antall gyldige differanser (***n***)
- Rangsummer (***W+, W-***)
- Teststatistikk ***W***
- P-verdi (eksakt eller z-approksimert)
- Effektstørrelse ***r***

Avvik fra SPSS?

- SPSS gir ofte begge rangsummer og velger statistikk basert på minste av dem.
- Verktøyet følger samme prinsipp og gir konsistente resultater.
- For større datasett brukes z-transformasjon og normalapproksimasjon, i tråd med SPSS sine avanserte prosedyrer.

Friedman-test

Friedman-testen undersøker om det er systematiske forskjeller mellom tre eller flere sammenkoblede målinger på samme enhet (f.eks. deltakere). Den er en ikke-parametrisk analog til én-veis ANOVA med repeterte målinger, og krever ikke normalfordelte data.

Når brukes Friedman-test?

- Når du har tre eller flere målinger på samme enhet (f.eks. pre / post / follow-up).
- Når du ikke kan anta normalfordeling.

Beregningsmetode

Rangering av data

For hver deltaker rangeres målingene:

- Laveste verdi får rang 1, nest laveste rang 2, osv.
- Dersom verdier er like (ties), tildeles midtverdien av rangeringsplassene.

Eksempel: verdiene [5, 7, 5] får rangene [1.5, 3, 1.5].

---

2. Beregning av teststatistikk

Formelen for Friedman-testens Chi^2-statistikk:

$$\chi^2 = \frac{12}{nk(k+1)} \sum_{j=1}^{k} R_j^2 - 3n(k+1)$$

n = antall deltakere

k = antall målinger (betingelser)

R_j = sum av rangene for hver måling/betingelse j

3. Beregning av p-verdi

$$p = 1 - \mathrm{CDF}_{\chi^2}(\chi^2, \mathrm{df} = k - 1)$$

P-verdien beregnes ved å sammenligne den observerte χ^2-verdien mot en kjikvadratfordeling med $k-1$ frihetsgrader de CDF er kumulativ fordelingsfunksjon.

4. Effektstørrelse: Kendall's W

$$W = \frac{\chi^2}{n(k-1)}$$

Friedman-testen gir også en mål på effektstørrelse, Kendall's W.

W = 0 → ingen enighet

W = 1 → perfekt enighet

Hva rapporterer verktøyet?

- Antall gyldige deltakere (***n***)
- Chi^2-verdi χ^2
- p-verdi
- Effektstørrelse (Kendall's W)
- Varsler om ulik skala/spredning mellom målingene
- Plot av gjennomsnittlige ranger for hver måling
- Automatisk post-hoc analyse hvis hovedtesten er signifikant

---

Post-hoc-analyse (Dunn's test)

$$z = \frac{|\bar{R}_i - \bar{R}_j|}{\sqrt{\frac{k(k+1)}{6n}}}$$

Ved signifikant hovedtest kjører verktøyet automatisk parvise sammenligninger mellom alle målinger, justert for multiple tester:

Teststatistikk for sammenligning mellom to målinger:

Hvor $\bar{R}i, \bar{R}j$ er gjennomsnittlige ranger for måling ***i*** og ***j***

- P-verdien justeres for antall sammenligninger (Bonferroni-korreksjon).

Mann-Whitney U-test

Mann-Whitney U-test (også kalt Wilcoxon rank-sum test) undersøker om to uavhengige grupper har forskjellig fordeling. Det er en ikke-parametrisk alternativ til uavhengig t-test.

Når brukes Mann-Whitney U?

- Når du sammenligner to uavhengige grupper (f.eks. gruppe A vs. gruppe B).
- Når data ikke er normalfordelt, eller måleskalaen er ordinal.
- Når du ønsker en test basert på rangering, ikke råverdier.

Beregningsmetode

1. Kombinerer og rangerer data

- Slår sammen alle observasjoner fra begge grupper.
- Rangerer alle verdier samlet (lavest verdi = rang 1 osv.).
- Ved "ties" (like verdier) tildeles gjennomsnittsrang.

---

2. U-verdien beregnes

Beregning av U for gruppe A hvor:

$$U_A = R_A - \frac{n_A(n_A+1)}{2}$$

- R_A = sum av rangene for gruppe A
- n_A = antall deltakere i gruppe A

3. Beregner z-verdi (for store utvalg)

For store utvalg (n > 20), brukes normalfordeling:

$$z = \frac{U - \mu_U}{\sigma_U}$$

Der forventet middelverdi =

$$\mu_U = \frac{n_A n_B}{2}$$

Og standardavvik =

$$\sigma_U = \sqrt{\frac{n_A n_B (n_A + n_B + 1)}{12}}$$

4. Beregner p-verdi

P-verdien finnes ved å bruke den standardiserte normalfordelingen hvor **Φ** er den kumulative fordelingsfunksjonen for normalfordelingen.

$$p = 2 \times (1 - \Phi(|z|))$$

5. Effektstørrelse (r)

Effektstørrelsen beregnes slik:

$$r = \frac{z}{\sqrt{n_A + n_B}}$$

Hva rapporterer verktøyet?

- Gjennomsnitt og standardavvik (SD) for hver gruppe
- U-verdi (indirekte via z)
- z-verdi
- p-verdi (two-tailed)
- Effektstørrelse ***r***
- Beskjed om resultatet er signifikant eller ikke

Kruskal-Wallis Test

Kruskal-Wallis-testen sammenligner tre eller flere uavhengige grupper for å undersøke om de kommer fra samme fordeling. Den er en ikke-parametrisk analog til én-veis ANOVA.

Når brukes Kruskal-Wallis?

- Når du har **tre eller flere uavhengige grupper**.
- Når dataene **ikke oppfyller kravene til normalfordeling**.
- Når data er **ordinal** eller har mange ekstreme verdier.

Beregningsmetode

1. Kombiner og ranger data

- Slå sammen alle observasjoner fra alle grupper.
- Ranger alle verdier samlet (laveste = rang 1, osv.).
- Ved like verdier (ties) gis gjennomsnittsrang.

2. Summer rangene for hver gruppe

For hver gruppe g:

- Summer alle rangene: R_g = sum av rangene i gruppe g

3. Beregn teststatistikk H

H-verdien beregnes slik:

$$H = \frac{12}{N(N+1)} \sum_{g=1}^{k} \frac{R_g^2}{n_g} - 3(N+1)$$

hvor:

- N = totalt antall observasjoner
- k = antall grupper
- n_g = antall observasjoner i gruppe g
- R_g = sum av ranger i gruppe g

4. Beregn p-verdi

P-verdien beregnes ved å sammenligne H med en kjikvadratfordeling (der df = antall grupper minus 1).

$$p = 1 - \mathrm{CDF}_{\chi^2}(H, \mathrm{df} = k - 1)$$

5. Effektstørrelse: ε^2 (epsilon squared)

Effektstørrelse for Kruskal-Wallis beregnes slik:

$$\varepsilon^2 = \frac{H}{N - 1}$$

Hva rapporterer verktøyet?

- Antall grupper
- H-verdi
- p-verdi
- Effektstørrelse ε^2
- Varsel hvis mindre enn 3 grupper
- Automatisk post-hoc tester (hvis signifikant)

Post-hoc-analyse

Ved signifikant Kruskal-Wallis, sammenlignes par av grupper med:

$$r = \frac{z}{\sqrt{n_1 + n_2}}$$

- Mann-Whitney U-test for hver parvis kombinasjon
- P-verdier justeres med Bonferroni-korreksjon:
- Effektstørrelse for hver sammenligning:
