

OPLOSSINGEN TESTVRAGEN

UNIVARIATE STATISTIEK

1. Deviatiescores zijn

- De som van de gekwadrateerde afwijkingen tegenover het rekenkundig gemiddelde
- De som van de afwijkingen tegenover het rekenkundig gemiddelde

2. De mediaan is een robuuste parameter van centraliteit

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

3. De Index van Diversiteit kan geïnterpreteerd worden als de kans dat twee willekeurig gekozen onderzoekseenheden tot een verschillende categorie behoren

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

4. De variatieratio neemt de waarde van nul aan indien alle waarnemingen tot de modale categorie behoren

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

5. Een variabele van het metrisch niveau kan bestudeerd worden op ordinaal niveau

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

6. Een variabele van het ordinale niveau kan bestudeerd worden op het metrische niveau

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

7. De variatiebreedte is het verschil tussen de maximale waarde en de minimale waarde

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

8. De variatiebreedte is de som van de maximale en minimale waarde

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

9. Uitspraken extrapoleren van de steekproef naar de populatie doe je via

- de beschrijvende univariate statistiek
- de inferentiële statistiek

10. De steekproefstandaardafwijking wordt berekend op basis van de formule van de populatiestandaardafwijking maar in de noemer staat $n+1$

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

11. Een bar chart wordt gebruikt

- Voor variabelen op het metrische niveau
- Voor variabelen op het categorische niveau

12. Variabelen die op een histogram worden gepresenteerd zijn steeds ratio niveau

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

13. Een lijndiagram wordt gebruikt voor de

- Visuele voorstelling van metrische gegevens die niet in klassen zijn gegroepeerd (ruwe scores)
- Visuele voorstelling van metrische gegevens die wel in klassen zijn gegroepeerd

14. Een frequentiepolygoon wordt gebruikt voor de

- Visuele voorstelling van metrische gegevens die niet in klassen zijn gegroepeerd (ruwe scores)
- Visuele voorstelling van metrische gegevens die wel in klassen zijn gegroepeerd

15. Een platykurtische variabele is

- Platter dan een normaal verdeelde variabele
- Scherper dan een normaal verdeelde variabele

16. Een box-plot kan worden gebruikt voor variabelen vanaf

- Het nominaal niveau
- Ordinaal niveau
- Interval niveau
- Ratio niveau

17. De mediaan komt overeen met

- Het vijftigste percentiel
- Het eenenvijftigste percentiel

18. Als we een kenmerk dat perfect normaal verdeeld is voorstellen via een box-plot, dan is de afstand tussen de mediaan en de hoogste waarde even groot als de afstand tussen de mediaan en de laagste waarde

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

19. Een variabele die rechtsscheef verdeeld is

- Heeft een langere staart naar rechts
- Heeft een langere staart naar links
- Heeft geen staart

20. De interkwartielafstand is een maat van

- Centraliteit
- Spreiding
- Vorm

21. Een frequentieverdeling kunnen we opvatten als een kansverdeling

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

22. Een onderzoekseenheid heeft een z-score van -2.20 voor het metrisch kenmerk “studieresultaat” .

- De onderzoekseenheid valt buiten de centrale 95% van de waarnemingen.
- De onderzoekseenheid doet het beduidend beter dan de gemiddelde onderzoekseenheid

23. Centreren wil zeggen dat men een kenmerk uitdrukt als een afwijking tegenover het rekenkundig gemiddelde

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

24. De frequentieverdeling (histogram) verandert vormelijk niet wanneer men standaardiseert

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

25. Het rekenkundig gemiddelde is een spreidingsmaat die gevoelig is voor uitschieters

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

26. Variabelen van het categorische niveau bevatten categorieën. Een onderzoekseenheid mag tegelijkertijd in twee categorieën van dezelfde variabele zitten

- Dit mag niet als we de regels van de statistiek volgen
- Dit mag wel, er zijn hier geen regels voor

27. Operationalisering betekent

- Dat we een kenmerk meetbaar maken
- Dat we een kenmerk van een conceptuele definitie voorzien

28. De populatievariantie is de som van de gekwadrateerde afwijkingen tegenover het rekenkundig gemiddelde, gedeeld door het steekproefeffectief

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

29. De variatieratio is een spreidingsmaat die we moeten gebruiken wanneer twee metrische kenmerken in een verschillende meeteenheid werden gemeten en men wil de spreiding vergelijken

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

30. De keuze voor een univariate parameter wordt ingegeven door

- De onderzoeksvraag
- Het meetniveau
- Beide

31. Stel: je bent activist die ijvert voor meer inkomensgelijkheid. Om de inkomensongelijkheid te demonstreren, kan je kiezen tussen een aantal parameters. Inkomen is zeer scheef verdeeld.

- Je presenteert vanuit je standpunt de mediaan
- Je presenteert vanuit je standpunt het rekenkundig gemiddelde

1. De Y-as wordt ook wel het ordinaat genoemd

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

2. De X-as wordt ook wel de abscis genoemd

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

3. Een scatterplot kan gebruikt worden om de relatie tussen een ratio-variabele en een interval-variabele te presenteren

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

4. Een scatterplot kan gebruikt worden om de relatie tussen een ordinale variabele en een interval-variabele te presenteren

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

5. Een spreidingsdiagram kan gebruikt worden om de relatie tussen een nominale en een ordinale variabele te presenteren

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

6. Het (x,y) coördinaat van het rekenkundig gemiddelde van x en het rekenkundig gemiddelde van y is het bivariate zwaartepunt van het spreidingsdiagram

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

7. De ongestandaardiseerde regressiecoëfficiënt (b_1) uit een bivariate regressieanalyse is een

- Symmetrische maat
- Asymmetrische maat

8. De bivariate gestandaardiseerde regressiecoëfficiënt uit de regressie van Y op X is gelijk aan de gestandaardiseerde covariantie tussen x en y

- Deze uitspraak is juist
- Deze uitspraak is fout

9. De regressie van Q op S wil zeggen dat we

- Q als afhankelijke variabele hebben en S als onafhankelijke variabele
- S als afhankelijke variabele hebben en Q als onafhankelijke variabele

10. Het percentageverschil is een symmetrische associatiemaat

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

11. Hierna vind je een tabel waarbij de relatie tussen twee nominale variabelen, X en Y wordt voorgesteld. We veronderstellen dat R de afhankelijke variabele is en dat S een causale invloed uitoefent op R.

| | Variabele S | | | Totaal |
|-------------|-------------|---|---|--------|
| Variabele R | A | B | E | |
| | C | D | F | |
| Totaal | G | H | I | |

Welke van de uitspraken is juist:

- E en F noemen we kolommarginale
- E en F noemen we rijmarginale
- A / G geeft de proportie van A
- A / C geeft de odds op A

12. Chi-kwadraat is een associatiemaat die zeer gevoelig is aan N

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

13. Als je N vermenigvuldigt met 2, dan wordt de waarde van chi-kwadraat twee keer zo groot

- Deze uitspraak is juist
- Deze uitspraak is fout

14. Chi-kwadraat moet worden berekend op basis van de ruwe scores

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

15. Chi-kwadraat kan ook worden berekend op basis van de proporties

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

16. In een 2*2 tabel is de waarde van Phi gelijk aan de waarde van V

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

17. Bekijk volgende tabel

| Slachtofferschap afgelopen vijf jaar | Scholingsgraad | | |
|--------------------------------------|----------------|--------|------|
| | Laag | Midden | Hoog |
| Nooit | A | B | C |
| Een maal | D | E | F |
| Twee maal of meer | G | H | I |

Welke uitspraken over deze tabel zijn juist

- (E*F) is een consistent paar
- (F*H) is een inconsistent paar

- Gamma is gebaseerde op de verhouding tussen de consistente paren min de inconsistente paren en de consistente paren plus de inconsistente paren
- Gamma is een associatiemaat die monotoniciteit veronderstelt
- Gamma is een asymmetrische maat
- Gamma wordt gebruikt op nominaal niveau

18. Welke uitspraken zijn juist?

- Spearman's Rho is een rangcorrelatiecoëfficiënt
- Pearson's r is een rangcorrelatiecoëfficiënt
- Kendall's Tau- is een rangcorrelatiecoëfficiënt
- Rangcorrelaties zijn asymmetrische maten

19. Welke uitspraken over Pearson's r zijn juist?

- De bivariate correlatiecoëfficiënt van Pearson veronderstelt lineariteit (zoniet kan Pearson's r leiden tot verkeerde conclusies)
- De bivariate correlatiecoëfficiënt is een gestandaardiseerde covariantie
- De bivariate correlatiecoëfficiënt wordt berekend uit niet gekwadrateerde deviatiescores

20. De bivariate correlatiecoëfficiënt tussen ouderlijk toezicht en criminaliteit en y is negatief en de gestandaardiseerde rco uit de regressie van criminaliteit op toezicht bedraagt -0.35.

- Daaruit volgt dat toezicht een remmend effect heeft op criminaliteit
- Daaruit volgt dat toezicht nefast is voor de criminaliteit want meer toezicht, resulteert in meer criminaliteit
- Daaruit volgt dat meer toezicht ongerelateerd is aan criminaliteit

21. De covariatie is

- De kruisproductensom
- De som van de deviatiescores van x + de deviatiescores van y

22. Hieronder volgen een aantal uitspraken over de lineaire regressieanalyse

- De lineaire regressie is een asymmetrische techniek
- De lineaire regressie (het basismodel) is gebaseerd op de kleinste kwadratenoplossing
- De lineaire regressie is niet geschikt voor kwadratische curvilineaire relaties

23. Het residu is

- Het verschil tussen geobserveerde en verwachte waarde
- Het verschil tussen verwacht en geobserveerde waarde

24. De som van de gekwadrateerde residuen bedraagt nul

- Deze uitspraak is juist
- Deze uitspraak is fout

25. De variatie in Y kan ook uitgedrukt worden als de som van de regressie sum of square minus de residual sum of square

- Deze uitspraak is juist
- Deze uitspraak is fout

26. Als de punten uit een puntenwolk heel dicht bij de regressierechte liggen, dan kunnen we vermoeden dat

- De model fit zeer hoog zal zijn
- De model fit zeer laag zal zijn

27. Het is perfect mogelijk dat twee regressiecoëfficiënten een zelfde beta-waarde hebben, maar een verschillende aliënatiecoëfficiënt hebben

- Deze uitspraak is juist
- Deze uitspraak is fout

28. De variantieanalyse wordt gebruikt wanneer

- De onafhankelijke variabele nominaal is en de afhankelijke variabele metrisch
- De afhankelijke variabele nominaal is en de onafhankelijke variabele metrisch

29. Eta-kwadraat komt qua interpretatie overeen met een determinatiecoëfficiënt

- Deze uitspraak is juist
- Deze uitspraak is fout

30. De toets van de significantie van de determinatiecoëfficiënt gebeurt aan de hand van

- De F-toets
- De T-toets

31. De toets van het intercept en de rico gebeurt aan de hand van de

- De F-toets
- De T-toets
- De Z-toets

32. Intervalschatting is een centraal element uit de

- Inferentiële statistiek
- Descriptieve statistiek

33. De steekproevenverdeling van X is

- De verdeling van X
- De verdeling van alle gemiddelde waarden voor X in een reeks van steekproeven met gelijke omvang
- De verdeling van alle gemiddelde waarden voor X in een reeks van steekproeven met ongelijke omvang
- Geen van de voorgaande beweringen is juist

**34. Een betrouwbaarheidsinterval van met een alfa (kans op vergissing) van 10 %
= Z-score van 90**

- Komt overeen met een z-score van 1.645
- Komt overeen met een z-score van 1.960
- Komt overeen met een z-score van 2.576

35. Ik doe criminologisch onderzoek naar belastingontduiking en stel vast dat het gemiddeld aantal veroordelingen onder belastingontduikers (mean = 3) lager is dan onder inbrekers (mean = 5). De betrouwbaarheidsintervallen van de beide gemiddeldes blijken elkaar te overlappen.

Hieruit besluit ik dat

- De gemiddeldes niet significant van elkaar verschillen
- De gemiddeldes wel significant van elkaar verschillen

36. Een type I fout betekent :

- De nulhypothese is juist en ik verwerp ze foutief
- De nulhypothese is verkeerd en ik behoud ze foutief

37. Een type II fout betekent :

- De nulhypothese is juist en ik verwerp ze foutief
- De nulhypothese is verkeerd en ik behoud ze foutief

38. De power van een test berekenen is belangrijk want hierdoor houdt men rekening met zowel type I als type II fouten

- Deze uitspraak is juist
- Deze uitspraak is fout

39. De tussengroepsvariantie is de tussengroepsvariantie gedeeld door het aantal vrijheidsgraden tussen groepen

- Deze uitspraak is juist
- Deze uitspraak is fout

40. Als de tussengroepsvariantie groter is dan de binnengroepsvariantie, en de F-toets geeft een waarde die hoger is dan de kritische waarde, is het verband tussen x en y significant

- Deze uitspraak is juist
- Deze uitspraak is fout

41. Als de tussengroepsvariantie groter is dan de binnengroepsvariantie, en de F-toets geeft een waarde die lager is dan de kritische waarde, is het verband tussen x en y significant

- Deze uitspraak is juist
- Deze uitspraak is fout

| |
|--------------------------------|
| MULTIVARIATE STATISTIEK |
|--------------------------------|

1. De meervoudige regressieanalyse met X1 en X2 als onafhankelijke variabelen geeft hetzelfde resultaat als twee afzonderlijke bivariate regressieanalyses. Klopt dat?

- Dit klopt enkel in de situatie waarbij $r(x_1, x_2) = 0$.
- Dit klopt nooit

2. R (hoofdletter) staat in de output van een multiële regressie voor de correlatie tussen Y en de verwachte waarde voor Y op basis van de onafhankelijke variabelen

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

3. Een meervoudige multivariate regressieanalyse is een

- Regressieanalyse met meerdere onafhankelijke variabelen en meerdere afhankelijke variabelen
- Regressieanalyse met meerdere onafhankelijke variabelen en slechts een afhankelijke variabele
- Regressieanalyse met meerdere afhankelijke variabele en een onafhankelijke variabele

4. Heteroscedasticiteit wil zeggen dat

- De waarde van de residuele termen toenemen naarmate X1 toeneemt
- De waarde van de residuele termen gelijk blijft, naarmate X1 toeneemt

5. Additiviteit betekent dat X_1 en x_2 onafhankelijke effecten hebben, dit wil zeggen: ze dragen elk bij tot de verklaring van Y

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

6. Lineariteit wil zeggen dat Y als lineaire functie van X kan worden uitgedrukt

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

7. Curvilineariteit wil zeggen dat het effect van X_1 op Y toeneemt of afneemt naargelang X_1 toeneemt.

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

8. Een uitbijter kan de regressielijn ernstig beïnvloeden

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

9. Interactie betekent dat het effect van X_1 op Y conditioneel is op X_2

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

10. Een mediatorvariabele is een variabele die het effect van een exogene variabele op een afhankelijke variabele medieert (dwz dat deze variabele het effect van de exogene variabele wegverklaart en dat het effect van de exogene variabele op Y via de mediator verloopt)

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

11. Een totaal effect is in een pad-analyse gelijk aan de rechtstreekse effecten min de onrechtstreekse effecten

- Deze uitspraak is juist
- Deze uitspraak is verkeerd

