

Sociale statistiek: enkele examenvragen van 2009-2010

1. (3 punten) Vul het gepaste woord/woorden in op de puntjes of doorstreep het woord dat niet past:

De covariantie is een maat van voor twee
..... variabelen. Ze is gebaseerd op het samen optreden van de
..... De absolute waarde hangt *wel / niet* af van de
meeteenheid en is daarom *zeer / weinig* informatief. Het teken van de covariantie
geeft informatie over de van het verband. Er wordt geen
onderscheid gemaakt tussen en De
covariantie wordt groter en negatief naarmate er zich meer punten in segment en
segment bevinden. Er is geen vaste

Een groot bedrijf vraagt zich af welke factoren bepalend zijn voor het al dan niet geven van een loonsverhoging aan zijn werknemers. Ze stellen een onderzoeker aan. Deze neemt van elke afdeling willekeurig 2 werknemers en noteert voor ieder de toename in loon van 15 jaar geleden tot nu. Verder kijkt hij op welk niveau de werknemers zich bevinden binnen het bedrijf (arbeiders, ploegbazen, afdelingsverantwoordelijken, directie).

Uit de resultaten blijkt dat de werknemers op directieniveau een grotere loonstoenamen kenden dan de werknemers op arbeidersniveau. De onderzoeker besluit hieruit dat het al dan niet krijgen van een loonsverhoging bepaald wordt door het werkniveau.

Een andere onderzoeker heeft hier zo zijn bedenkingen bij en stelt dat het geslacht van de werknemers dit effect bepaalt. In de bedrijfsvoorschriften wordt vooropgesteld dat er op elk niveau 50% mannen en 50% vrouwen moeten zijn, zoals in de globale populatie. In de steekproef is dit echter niet zo. 80% van de werknemers op het directieniveau zijn mannen. Op het arbeidersniveau zijn er slechts 30% mannen.

De eerste onderzoeker gebruikt als variabele het werkniveau.

Dit is een (*specifiek meetniveau*) variabele. De

..... variabele is het loonsverschil tussen het huidige loon en dat van

15 jaar geleden. Dit is een (*specifiek meetniveau*) variabele. De

andere onderzoeker ziet het geslacht als een variabele. Dit is

een (*specifiek meetniveau*) variabele. 50% is een

..... terwijl 30% en 80% zijn.

OPLOSSING

De covariantie is een maat van SAMENHANG voor twee KWANTITATIEVE / METRISCHE variabelen. Ze is gebaseerd op het samen optreden van de DEVIATIESCORES / AFWIJINGEN TOV HET GEMIDDELDE. De absolute waarde hangt *wel* / ~~niet~~ af van de meeteenheid en is daarom ~~zeer~~ / *weinig* informatief. Het teken van de covariantie geeft informatie over de RICHTING van het verband. Er wordt geen onderscheid gemaakt tussen OV / EXPLANATORY VAR / VERKLARENDE VAR en AV / OUTCOME VAR / RESPONS VAR. De covariantie wordt groter en negatief naarmate er zich meer punten in segment II en segment IV bevinden. Er is geen vaste BOVEN- EN ONDERGRENS.

De eerste onderzoeker gebruikt als ONAFH / EXPLANATORY / VERKLARENDE variabele het werkniveau. Dit is een ORDINALE (*specifiek meetniveau*) variabele. De AFH / OUTCOME / RESPONS variabele is het loonsverschil tussen het huidige loon en dat van 15 jaar geleden. Dit is een KWANTITATIEVE / METRISCHE (*specifiek meetniveau*) variabele. De andere onderzoeker ziet het geslacht als een LURKING / LATENTE / VERBORGEN variabele. Dit is een NOMINALE (*specifiek meetniveau*) variabele. 50% is een PARAMETER terwijl 30% en 80% STATISTIEKEN zijn.

2. (2 punten) Starttermen Wiskunde: Gegeven onderstaande 4x4 kruistabel van de kenmerken X en Y.

| | | Y | | | | |
|---|---|----|----|----|----|-----|
| | | 1 | 2 | 3 | 4 | |
| X | 1 | 19 | 5 | 7 | 21 | 52 |
| | 2 | 23 | 25 | 11 | 17 | 76 |
| | 3 | 31 | 34 | 29 | 27 | 121 |
| | 4 | 11 | 13 | 17 | 19 | 60 |
| | | 84 | 77 | 64 | 84 | |

Bereken $\sum_{i=2}^3 \sum_{j=3}^4 3 * f_{ij}^2 + 5 * f_{..} =$

Berekening:

$$\sum_{i=2}^3 \sum_{j=3}^4 3 * f_{ij}^2 + 5 * f_{..} = (3 * 11^2 + 3 * 17^2 + 3 * 29^2 + 3 * 27^2) + 5 * 309 = 7485$$

Antwoord:

3. (3,5 punten) Een OCMW heeft 5 bedragen die uitgekeerd worden als leefloon. De relatieve frequentieverdeling van de uitkeringen van dit OCMW is weergegeven in onderstaande tabel.

| Uitkering | f' |
|-----------|----|
| 750 | 10 |
| 900 | 10 |
| 1000 | 40 |
| 1150 | 25 |
| 1250 | 15 |

- a) Bereken de variantie van de uitkering.
 b) Men beslist om de uitkering van 750 euro op te trekken met 10 procent en dit te financieren door de hoogste uitkering te doen dalen. Met hoeveel procent moet men de hoogste uitkering doen dalen om de verhoging van de laagste uitkering bekostigd te krijgen?

Berekening:

a)

$$\bar{x} = \sum_{i=1}^m p_i * x_i = 0.10 * 750 + 0.10 * 900 + 0.40 * 1000 + 0.25 * 1150 + 0.15 * 1250 = 1040$$

$$s^2_x = \sum_{i=1}^m p_i * (x_i - \bar{x})^2 =$$

$$0.10 * (750 - 1040)^2 + 0.10 * (900 - 1040)^2 + 0.40 * (1000 - 1040)^2 + 0.25 * (1150 - 1040)^2 + 0.15 * (1250 - 1040)^2 = 20650$$

b) aantal observaties = n

10% van 750 is 75

$$75 * n * 0.10 = 7.5n$$

Er moet 7.5n euro bij de bovenste categorie bij, dus moet er ook zo veel van de onderste categorie af:

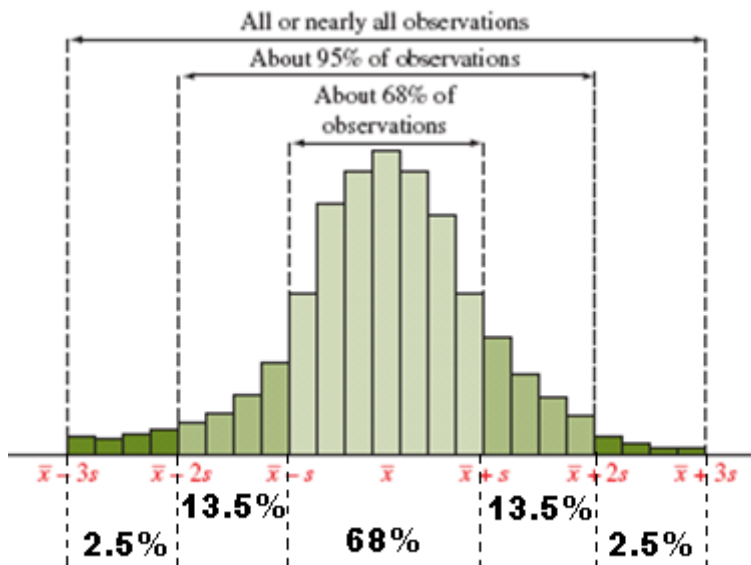
$$\text{Resterend bedrag in de onderste categorie: } 1250 * n * 0.15 - 7.5 * n = 180n$$

Dit bedrag moet verdeeld worden over $0.15 * n$ gezinnen. Er is dus voor elk gezin $180n / 0.15n = 1200$ euro. Er was 1250 euro. Na de herverdeling is er nog maar 1200 euro. Dit is een daling van 50 euro of van $50 / 1250 * 100 = 4\%$

4. (3 punten) Gegeven: een klokvormige verdeling. 16% van de observaties heeft een waarde kleiner dan of gelijk aan 17. De waarde 26 komt overeen met een z-score van 2. Bepaal het gemiddelde en de variantie van deze verdeling.

Berekening:

Tekening:



$$\bar{x} - s_x = 17$$

$$2 = \frac{26 - \bar{x}}{s_x}$$

$$z_i = \frac{26 - \bar{x}}{s_x} = 2 \Leftrightarrow 26 - \bar{x} = 2s_x \Leftrightarrow \frac{26 - \bar{x}}{2} = s_x$$

$$\bar{x} - s_x = 17 \Leftrightarrow \bar{x} - \frac{26 - \bar{x}}{2} = 17 \Leftrightarrow \bar{x} + \frac{\bar{x}}{2} = 17 + 13 \Leftrightarrow \frac{3\bar{x}}{2} = 30 \Leftrightarrow \bar{x} = \frac{30 * 2}{3} = 20$$

$$\bar{x} - s_x = 17 \Leftrightarrow 20 - s_x = 17 \Leftrightarrow s_x = 3 \Leftrightarrow s^2_x = 9$$

| |
|---|
| <p>Antwoord:</p> <p style="text-align: center;">Gemiddelde: Variantie:</p> |
|---|