Elementen van kanstheorie en statistiek

Elementen van kanstheorie en statistiek	
Richting	Wiskunde
Jaar	1BWIS

Bespreking

Dit vak wordt sinds 2013 gegeven door professor Volders en het is grotendeels herhaling van de kansrekening en statistiek van het middelbaar, maar uiteraard wordt er wel iets dieper op in gegaan. Voornamelijk zal je een hele boel extra verdelingen leren.

Sinds 2014 wordt vak door professor Guilaume gegeven. Gezien haar lessen zal ze tegen volgend jaar, vermoedelijk, een eigen, uitgebreidere, cursus voorzien.

Theorie

Er is een cursus voor dit vak die alle geziene leerstof omvat. De theorielessen houden min of meer in dat de cursus op het bord geschreven wordt, met eventueel extra aanvullingen zoals een niet gegeven bewijs.

Oefeningen

De oefeningen worden gegeven door Valérie De Witte en zijn in tegenstelling tot de theorie heel verhelderend. Ze sluiten nauw aan bij de cursus en zorgen ervoor dat de iets moeilijkere stukken theorie makkelijker te begrijpen zijn. Het is dus aan te raden om naar de oefeningenlessen te gaan, zeker als tijdens de theorie niet alles duidelijk was.

Puntenverdeling en examen

Het examen wordt afgelegd in één keer. Je krijgt wel apart de theorie, en daarna pas de oefeningen. Bij de theorie is een rekenmachine en formularium NIET toegestaan, bij de oefeningen is dit wel toegelaten (rekenmachine mag niet grafisch zijn). De meeste vragen zijn oefeningen of toepassingen op de theorie. Uiteraard moet je ook wel voorbereid zijn op enkele pure theorievragen, maar die nemen zeker niet het grootste deel van het examen in.

Theorie en oefeningen tellen beide mee voor 50% van het totaal.

Examenvragen

Juni 2012 (prof Verhasselt)

1. Juist of fout (+ toon aan / leg uit):
    - Stel X een strikt positieve stochastische veranderlijke. Dan geldt dat dat de interkwartielafstand van sqrt(X) gelijk is aan de wortel van de interkwartielafstand van X.
    - Om normaal verdeelde gegevens op een normale QQ-plot voor te stellen, moet je ze eerst standaardiseren.
    - Als $X\sim t4X\sim t4$ en $Y\sim t6Y\sim t6$ dan geldt dat de kans dat X kleiner is dan -3.5, kleiner is dan de kans dat Y groter is dan 3.5.
    - Om een bepaalde parameter te onderzoeken van een of andere verdeling voert Jantje 100 metingen uit en vindt een 95%-betrouwbaarheidsinterval voor die parameter. Liesje doet hetzelfde, maar dan met 200 metingen. Het betrouwbaarheidsinterval van Liesje is kleiner dan dat van Jantje.
2. Bewijs:
    - Lemma van Slutsky
    - Als $Xn\to DXXn\to DX$ en $Yn\to PcYn\to Pc$ (c reëel), dan geldt
      $(Xn+Yn)\to D(X+c)$

$$(Xn+Yn)\to D(X+c)$$

3. Op een parking staan n auto's, en in één van de auto's zit een bom verborgen. De kans dat de bom in auto i zit is gelijk aan pipi (i = 1, ..., n). Als de bom in auto i zit, is de kans dat ze bij onderzoek van de auto gevonden wordt gelijk aan aiai. Zoek de waarde van BijBij, die de kans voorstelt dat de bom in auto j zit als gegeven is dat de bom niet in auto i gevonden werd.
4. Een paar eeuwen geleden zat ook de wiskundige Paul de Méré met een paradox opgezadeld. Bij een spel waar drie dobbelstenen gerold moesten worden, stelde hij experimenteel vast dat de som van de ogen vaker 11 was dan 12. Theoretisch gezien was dit volgens hem onmogelijk, men had immers zes mogelijkheden op 11 en zes op 12. Deze gebeurtenissen moesten dus gelijke kans hebben! Waar zit de fout in zijn redenering en wat zijn de exacte kansen?
5. *Gegeven: de Gumbel-verdeling met gegeven dichtheidsfunctie, variantie, gemiddelde en Eulerconstante.*
    - Bereken de verdelingsfunctie
    - *Transformatie van een Gumbel-verdeelde X naar een Y = AX + B*
    - Is Tn=1n·∑ni=1XiTn=1n·∑i=1nXi een onvertekende schatter voor thèta? Zo niet: geef een onvertekende schatter.
    - Is je gevonden schatter consistent?
6. Het examen statistiek staat op 20 punten. Elke student kan dus 0, 1, 2, ..., 19 of 20 punten halen. Stel M de hoogste behaalde score in de klas, en stel dat er n studenten in de klas zitten. Bereken:
    - De kans dat M = 0.
    - De kans dat M = 1.
    - De kans dat M = k (voor k = 1, 2, ..., 20)
7. Na uitvoerig onderzoek vond men het volgende 95%-betrouwbaarheidsinterval voor het zoutgehale in de Rode Zee: [38.9, 40.9] in de veronderstelling dat dat normaal verdeeld is met variantie s2s2. Geef het 99%-betrouwbaarheidsinterval.

Augustus 2012

1. Juist of fout (+ toon aan / leg uit):
    - Van een bepaald stochastische veranderlijk doet men een steekproef bij 250 proefpersonen. Men maakt hiervan een 90% betrouwbaarheidsinterval. Stel dat men nog een steekproef doet bij 250 andere proefpersonen, de kans dat het steekproefgemiddelde bij deze 250 proefpersonen binnen het 90% betrouwbaarheidsinterval zit, is 90%.
    - De centrale limietstelling impliceert dat hoe groter het aantal steekproefelementen is , hoe meer dat de gestandaardiseerde som gaat lijken op de normale verdeling
    - Stel men heeft een verdeling van een stochastisch veranderlijke. Men maakt een QQ-plot waarbij men deze vergelijkt met de theoretische kwantielen van de exponentiële verdeling. Als de QQ-plot een exponentiële functie voorstelt, wil dat zeggen dat de stochastische veranderlijke verdeeld is volgens een exponentiële verdeling.
    - X 1 2 3 4

P 0.2 0.5 0.1 0.2 en stel $Y= 2X^2 + 6$. a) E(X)= 3 b) Var(X)= 0.34 c) E(Y)= 18.6 d) Var(Y)= 4.3

1. a) Geef de definitie van de centrale limietstelling en geef hiervan ook een voorbeeld

b) Definieer de ongelijkheid van Chebyshev en geef het bewijs voor het discrete geval.

1. Stel in een parochie van een welbepaald dorp worden er lotjes verkocht ten voordele van de parochie. Er zijn in totaal 100 lotjes en 4 ervan geven recht op een hoofdprijs en 10 geven recht op een troostprijs. De pastoor geeft het goede voorbeeld en koopt 5 lotjes.
    - a) wat is de kans dat hij tenminste 2 prijzen heeft?
    - b) Wat is de kans dat hij precies 1 hoofdprijs haalt en 1 troostprijs haalt ?
    - c) wat is de kans dat hij precies 1 hoofdprijs haalt maar geen troostprijs haalt?

1. Stel een luchtvaartorganisatie heeft allerlei motorvliegtuigjes. Stel dat de kans dat een motor van een vliegtuigje kapot is, pp is. Een vliegtuig is kapot als minder dan de helft van de motoren defect zijn. Voor welke pp is de kans dat een 4 motorig vliegtuig defect is groter dan de kans dat een 2-motorig vliegtuigje defect is?

Juni 2013 (prof Volders)

Theorie

Zijn volgende beweringen juist of fout? Indien juist, bewijs. Indien fout, weerleg met een tegenvoorbeeld.

1. Σnajbn−j=(a+b)nΣnajbn−j=(a+b)n
2. P(A∪B)=P(A)+P(B)+P(A∩B)P(A∪B)=P(A)+P(B)+P(A∩B)
3. Var(−X)=−Var(X)Var(−X)=−Var(X)
4. Var(X+Y)=Var(X)+Var(Y)Var(X+Y)=Var(X)+Var(Y)
5. Var(X2)<E(X4)Var(X2)<E(X4)

Volgende vragen werden ook zo gesteld, maar moesten in feite gewoon bewezen worden:

1. Ongelijkheid van Chebyshev
2. Momentgenererende functie van de chi-kwadraat-verdeling.
3. Variantie Poisson-verdeling

Definities:

1. mediaan
2. discrete stochastische veranderlijke

Oefeningen

1. Gegeven is dat de verdeling normaal verdeeld is. Van die verdeling krijg je het gemiddelde en P(X<90P(X<90. Hierover krijg je enkele standaard vragen over de normale verdeling: bereken de variantie en standaardafwijking. Wat is de kans dat X > ...
2. Bij de tweede vraag werd alle informatie over de normale verdeling gegeven en moest je de kans berekenen dat je tussen twee grenzen zit.
3. Iemand telt de auto's die langs haar huis voorbijrijden. Ze telde er 2500 in één uur. Wat is de kans dat ze de volgende minuut er minder dan 20 telt?
4. In een gevangenis zitten drie gevangenen. Er wordt beslist dat twee van de gevangenen geëxecuteerd worden. Wie blijft echter geheim voor de gevangenen. Een van de gevangenen (A) vraagt toch aan de cipier om 1 iemand te geven. De cipier denkt even na en zegt dat B geëxecuteerd zal worden, omdat hij denkt dat de kans onveranderd blijft. A bedankt de cipier en is opgelucht, zijn overlevingskans is gestegen zo denkt hij. Wie heeft er gelijk? De cipier of de gevangene en waarom?
5. In de Hema wordt er een actie gehouden waarbij je bij elke aankoop een letter van HEMA krijgt. Na hoeveel keren verwacht men dat men dit woord heeft?
6. Een oefeningen over schatters en betrouwbaarheidsintervallen. (Niet 100% zoals in de cursus).

Juni 2014 (prof Guillaume)

Theorie

1. Beschouw een experiment met een dobbelsteen. Is de verzameling {ϕ,{1},{2},{2,3,4,5,6},{1,3,4,5,6},Ω}{ϕ,{1},{2},{2,3,4,5,6},{1,3,4,5,6},Ω} een σσ - algebra?
2. Stel dat je een bepaald experiment blijft herhalen tot je eerste succes. Zij X dan het aantal mislukkingen voor je eerste succes met verwachtingswaarde 6.5 en Y het aantal pogingen tot en met het eerste succes. Wat zijn verwachtingswaarde en variantie van Y?
3. Beschouw de dichtheidsfunctie
    - fX(x)=⎡⎣⎢⎢⎢⎢⎢⎢0(x≤0)k.x2(0≤x≤2)b.x+c(2≤x≤a)0(a≤x)fX(x)={0(x≤0)k.x2(0≤x≤2)b.x+c(2≤x≤a)0(a≤x)
    - met a, b, c en k constant. Bepaal dan a zodat de mediaan gelijk is aan de modus (uit een bijhorende grafiek was af te lezen dat de modus 2 is).
4. Gegeven is volgende momentgenererende functie
    MX(t)=18.e−3.t+14.e−2.t+16+16.et+112.e2.t+524.e3.t

    MX(t)=18.e−3.t+14.e−2.t+16+16.et+112.e2.t+524.e3.t

    Bepaal P(X≤2)P(X≤2).
5. Beschouw de volgende drie intervallen
    I1=[x‾‾‾n−2,81.σn−−√;x‾‾‾n+2,42.σn−−√]I2=[x‾‾‾n−2,56.σn−−√;x‾‾‾n+2,56.σn−−√]I3=[−∞;x‾‾‾n+2,34.σn−−√]

    I1=[x¯n−2,81.σn;x¯n+2,42.σn]I2=[x¯n−2,56.σn;x¯n+2,56.σn]I3=[−∞;x¯n+2,34.σn]

Welke intervallen beschrijven een 99% BTI?

(a) Geen van de drie

(b) Enkel I1I1

(c) Enkel I2I2

(d) Enkel I3I3

(e) I1I1 en I2I2

(f) I1I1 en I3I3

(g) I2I2 en I3I3

(e) I1,I2I1,I2 en I3I3

Oefeningen

1. Combinatoriek
    - Hoeveel getallen van 4 verschillende cijfers kan je maken met 1, 2, 4, 5 en 7 ?
    - Hoeveel van deze getallen beginnen met 5, 54 en 542 ?
    - Als je de getallen rangschikt van klein naar groot. Welk getal staat er dan op de 50ste plaats? Op welke plaats staat het getal 4521 ?
    - Hoeveel getallen zijn er waar 5, 4 en 2 in deze volgorde naast elkaar staan ?
    - Hoeveel getallen zijn er waar 5, 4 en 2 in willekeurige volgorde naast elkaar staan?
2. Op een middelbare school worden 's middags extra lessen wiskunde gegeven ter voorbereiding van het ingangsexamen geneeskunde. Van de leerlingen die aan het examen zullen deelnemen, volgt 80% deze lessen. De kans dat een leerling die deelneemt aan de lessen slaagt voor het examen is 55%. De kans dat een leerling die niet deelneemt aan de lessen niet slaagt voor het examen is 65%. Een studente nam deel aan het examen en slaagde. Hoe groot is de kans dat ze de lessen volgde?
3. Het gewicht van een pilletje is normaal verdeeld met een gemiddelde van 3 gram en een standaarddeviatie van 0,2 gram. Een patiënt moet van de dokter 10 dagen lang 10 pillen slikken. Hij koopt een doos van 100 pillen. De doos zelf weegt 27 gram en de bijsluiter 19 gram.
    - Bereken de kans dat de volledige doos (i.e. doos met pillen en bijsluiter) meer dan 350 gram weegt.
    - Bereken de kans dat bij de eerste 10 pillen die hij slikt, er 4 pillen zijn waarvan het gewicht zich tussen 2,86 gram en 3,16 gram bevindt.
4. In een bepaalde regio lijdt 5% van de bevolking aan een vorm van suikerziekte. Om te bepalen wie deze ziekte heeft, worden er bloedstalen genomen. Het onderzoeken van bloedstalen is echter tijd- en geldrovend. Daarom besluit men om, in plaats van elk bloedstaal apart te onderzoeken, telkens k bloedstalen bij elkaar te doen en deze mix te testen. Als deze test negatief is heeft geen van de k personen suikerziekte en heeft men dus k - 1 onderzoeken uitgespaard. Als de test positief is, moet men de k mensen apart onderzoeken en heeft men 1 onderzoek verlies gedaan.
    - Bepaal verdeling en verwachtingswaarde van X, het aantal bloedstalen dat moet worden onderzocht per groep van k mensen.
    - Bepaal de verwachtingswaarde voor de winst die wordt gemaakt per groep van k mensen.
    - Bepaal de verwachtingswaarde voor de winst per persoon.
    - Bepaal k zodat er een minimum aan onderzoeken moet worden verricht.

Augustus 2014 (prof. Guillaume)

Theorie

1. Gegeven is een QQ-plot van een data set van waarnemingen. Door de plot te bekijken moet je bepalen welk van de volgende uitspraken geldig is:
    - de gebeurtenis standaard normaal verdeeld.
    - de gebeurtenis is normaal verdeeld met μ=1μ=1 en σ=2σ=2.
    - de gebeurtenis is normaal verdeeld met μ=2μ=2 en σ=1σ=1.
    - de gebeurtenis is niet normaal verdeeld.

2. Zij X een Geometrisch verdeelde stochastische variabele. $P(X \geq n+k \mid X \geq n) = P(X \geq k) n, k \geq 0$
    - Dit is waar.
    - Dit is niet waar.
    - Dit hangt af van n.
3. Zij X een stochastische variabele, dan $Var[X^n] \leq E[X^{2n}]$
    - Dit is waar.
    - Dit is niet waar.
    - Dit hangt af van n
    - Zij $M_X(t)$ de momentgenererende functie voor een stochastische variabele X. Is $2M_X(t)$ dan ook een momentgenererende functie voor een stochastische variabele?
4. Beschouw een 99% betrouwbaarheidsinterval [3.80, 4.30] voor het ongekende kental van een experiment. Welke uitspraak is/uitspraken zijn waar?
    - De kans dat de uitkomsten van de steekproef tussen 3.80 en 4.30 liggen 99%.
    - De kans dat het ongekende kental zich tussen 3.80 en 4.30 bevindt is 99%.
5. Beschouw het volgende experiment met een dobbelsteen

$$X(\omega) = \{+1\ \omega=\{1,3,5\}\ -1\ \omega=\{2,4,6\}\quad X_n(\omega) = \{+1\ \omega=\{1,2,3\}\ -1\ \omega=\{4,5,6\}$$

. Welke uitspraak is waar:
    - $X_1+X_2+...X_n$ convergeert in kans maar niet in verdeling.
    - $X_1+X_2+...X_n$ convergeert in kans en in verdeling.
    - $X_1+X_2+...X_n$ convergeert niet in kans en niet in verdeling.
    - $X_1+X_2+...X_n$ convergeert niet in kans maar wel in verdeling.

Oefeningen

1. Combinatoriek:
    - Hoeveel getallen van 3 cijfers kan je maken met de cijfers 0, 1, 2, 4, 5, 7, 9?
    - Hoeveel van deze getallen beginnen met 1? Hoeveel met 4?
    - Als je de getallen zou rangschikken naar waarde, welk getal staat dan op de 150ste plaats?
    - Op welke plaats staat dan het getal 275?
2. We testen een populatie op AIDS. Stel dat de kans dat een ziek persoon positief test 99% is, en dat deze kans gelijk is wanneer iemand die de ziekte niet heeft, negatief test.. In het algemeen is de kans om AIDS te hebben 0,6%. Hoe groot is dan de kans dat de test positief is op voorwaarde dat de persoon gezond is?
3. Bij het spel roulette kan je inzetten op 37 getallen: de oneven getallen zijn rood, de even getallen zijn zwart en nul is wit. Als je inzet op een getal en je wint, krijg je je inzet maal 36 terug. Stel dat de stochastische variabele gelijk is aan de winst of verlies die je maakt wanneer je 100 euro inzet:
    - Bepaal de verdeling van X en bereken de verwachtingswaarde.
    - Hoeveel winst/verlies verwacht je?
    - Bepaal de variantie en de standaarddeviatie.
4. Een bank accepteert rollen met muntstukken van 1 eurocent. Een rol bevat 50 centjes die de bank meteen terugbetaalt. Omdat het over een klein bedrag gaat, telt men dit vaak niet na. De kans dat er een cent te weinig in een rol zit is 30 procent, de kans dat er een cent te veel in de rol zit is 10 procent en de kans dat er exact 50 eurocent in zit is 60%. Stel X de stochastische variabele die staat voor het verlies bij een rol.
    - Hoe groot is de kans dat de bank 25 cent verlies maakt bij het aanvaarden van 100 rollen?
    - Hoeveel rollen met de bank minstens innen om netto verlies te maken?
    - Gegeven is een 95% [39.9,40.9] voor het gemiddelde van een steekproef met standaarddeviatie σ^2. Geef een 99% betrouwbaarheidsinterval voor dat gemiddelde.

Juni 2015

Theorie

1. Gegeven $P(A \cup B)$=0,9 en P($A \cup B^c$)=0,7. Aan hoeveel is $P(A)$ dan gelijk?
2. Is {$\emptyset, A, A^c, \omega$} een stam?
    - Waar.
    - Niet waar.

3. P(X>s+t|X>tX>s+t|X>t) = P(X>sX>s)?
    - Waar.
    - Niet waar.
    - Dit hangt af van t.

Oefeningen

1. Sinaasappelen hebben een gemiddeld gewicht van 50 g en een standaardafwijking van 10 g. Een zakje in de supermarkt heeft een draagkracht van max 1 kg. Hoeveel sinaasappelen mag men max in het zakje steken zodat de kans dat het zakje scheurt kleiner is dan 1 procent?
2. In een onderzoek naar de houding van Belgen tegen migranten stelt men dat op basis van een aselecte steekproef van 2200 Belgen, het percentage van de Belgische bevolking dat tegen het gemeentelijk stemrecht voor migranten is, geraamd kan worden op 63,02% à 66,24%. Hoe groot is de kans dat het percentage tegenstanders toch lager ligt dan dan het minimale percentage dat in het onderzoek bepaald werd?
3. Noem X de hoeveelheid verse aardbeien in kg dat de mensen tijdens de zomermaanden in een fruitwinkel willen kopen:
    - fX(x)={γ(100−x)(0≤x≤100)0(100<x)fX(x)={γ(100−x)(0≤x≤100)0(100<x)
    - αα is de prijs per kilogram waaraan de winkeleigenaar ze koopt, 2α2α verkoopt en de aardbeien die op het einde van de avond nog niet zijn verkocht verkoopt hij aan 0,5α0,5α. YsYsis de winst die de winkelier op het einde van de werkdag maakt indien hij smorgens verse aardbeien aankoopt(s≤100)(s≤100). Ys=gs(x)Ys=gs(x) met gsgseen functie die voor elke concrete waarde van X berekent wat de dagwinst is.
    - Zoek γγ zodat fXfX een dichtheidsfunctie is.
    - Zoek E[Y].
    - s zodat de dagwinst maximaal is.
4. Gegeven een rij van 5 cijfers die telkens gekozen worden uit(3,7,9).
    - Hoeveel verschillende rijen kunnen er gevormd worden met precies 4 keer een drie.
    - Een 9 op de eerste plaats.
    - De som van de cijfers minstens 35 is.