De multivariate normale verdeling

P.J.C. Spreij

Inhoudsopgave

1 Karakteristieke functies van univariate verdelingen

1.1 Definitie en eigenschappen

Eigenschappen van een (reëelwaardige) stochastische variabele kunnen afgeleid worden uit eigenschappen van bijvoorbeeld zijn dichtheid (als deze bestaat). Dit is echter lang niet altijd eenvoudig. Vandaar dat men in de kansrekening ook vaak werkt met transformaties van een verdeling (dichtheid). De transformatie die we in deze paragraaf bespreken is de overgang op de karakteristieke functie. Voordat we deze definiëren, merken we op dat er ook andere transformaties bestaan met analoge eigenschappen. We noemen in dit kader de begrippen kansgenererende functie (wordt o.m. gebruikt in de theorie van de Markovketens) en de Laplace transformatie, die veel in de wachtrijtheorie wordt gebruikt.
Karakteristieke functies zijn o.m. buitengewoon handig bij het bewijzen van bijvoorbeeld de centrale limietstelling. Een uitgebreide behandeling hiervan vindt plaats in een ander college.

Definitie 1.1 *Zij X een stochastische variabele. De complexwaardige functie ϕ met reëel argument van (de kansverdeling van) X is gedefinieerd door*

$$\phi(u) = \mathbb{E}e^{iuX} = \mathbb{E}\cos(uX) + i\mathbb{E}\sin(uX). \tag{1.1}$$

Indien in een bepaalde context sprake is van meer dan een stochastische variabele schrijven we om verwarring te voorkomen ook wel ϕ_X i.p.v. ϕ.

Als X een absoluut continu verdeelde stochast is met dichtheid f kunnen we dus schrijven $\phi(u) = \int e^{iux} f(x)dx = \int \cos(ux)f(x)dx + i\int \sin(ux)f(x)dx$. Is X daarentegen discreet verdeeld met waarden in $\{x_1, x_2, \ldots\}$, dan kunnen we schrijven $\phi(u) = \sum_{k=1}^{\infty} e^{iux_k} P(X = x_k)$. We geven gelijk een paar eigenschappen van karakteristieke functies.

Propositie 1.2 *Zij ϕ de karakteristieke functie van (de kansverdeling van) een stochastische variabele X. Dan heeft ϕ de volgende eigenschappen.*
(i) $|\phi(u)| \leq 1$ en $\phi(0) = 1$.
(ii) Als X symmetrisch om nul verdeeld is, is ook ϕ symmetrisch om nul. Bovendien is ϕ in dit geval reëelwaardig en $\phi(u) = \mathbb{E}\cos(uX)$.
(iii) Als $\mathbb{E}|X|^k < \infty$, dan geldt dat $\mathbb{E}X^k = (-i)^k \phi^{(k)}(0)$.

Bewijs. (i) Triviaal.
(ii) We bekijken alleen het geval waarin X absoluut continu verdeeld is met dichtheid f. Wegens de symmetrie geldt $f(x) = f(-x)$.
Maar dan is $sin(ux)f(x)$ een oneven functie in x en heeft dus integraal 0. En dus $\phi(-u) = \mathbb{E}\cos(-uX) = \mathbb{E}\cos(uX) = \phi(u)$. □

We geven nu wat voorbeelden.

Voorbeeld 1.3. Zij X alternatief verdeeld met parameter p. Dan is $\phi(u) = pe^{iu} + 1 - p$.

Voorbeeld 1.4. Zij X homogeen verdeeld op $[0,1]$, dan is $\phi_X(u) = \frac{e^{iu}-1}{iu}$ voor $u \neq 0$. Als Y homogeen verdeeld is op $[-1,1]$, dan is $\phi_Y(u) = \frac{\sin u}{u}$ voor $u \neq 0$.

Voorbeeld 1.5. De karakteristieke functie van de standaard normale verdeling wordt gegeven door $\phi(u) = e^{-\frac{1}{2}u^2}$. Dit beargumenteren we als volgt. Omdat de standaard normale verdeling symmetrisch is, kunnen we wegens propositie 1.2 schrijven

$$\phi(u) = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \cos(xu)dx.$$

We differentiëren deze vergelijking links en rechts naar u en verwisselen in het rechterlid integraal en afgeleide (waarom dit hier mag, laten we buiten beschouwing). Er komt

$$\phi'(u) = -\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} x \sin(xu)dx.$$

Op het rechterlid passen we partiële integratie toe, zodat

$$\phi'(u) = [\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \sin(xu)]_{-\infty}^{\infty} - u \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2} \cos(xu)dx.$$

Dus krijgen we de eenvoudige lineaire differentiaalvergelijking $\phi'(u) = -u\phi(u)$. Samen met de voorwaarde $\phi(0) = 1$ geeft dit de oplossing $\phi(u) = e^{-\frac{1}{2}u^2}$.

Veel bekende verdelingen zijn opgebouwd uit relatief eenvoudige bouwsteentjes. Een stochastische variabele met een Bernoulli verdeling bijvoorbeeld kunnen we zien als een som van onafhankelijke alternatief verdeelde stochasten. We zullen zien dat soortgelijke relaties ook gelden voor karakteristieke functies. Dit heeft het voordeel dat we niet steeds met allerlei ad hoc methoden karakteristieke functies hoeven uit te rekenen.

Propositie 1.6 *De volgende rekenregels gelden.*
(i) Als X een stochastische variabele is en $Y = aX + b$, dan is $\phi_Y(u) = e^{iub}\phi_X(au)$.
(ii) Als X en Y onafhankelijke stochastische variabelen zijn, dan heeft hun som $X + Y$ de karakteristieke functie gegeven door $\phi_{X+Y}(u) = \phi_X(u)\phi_Y(u)$.

Bewijs. (i) Volgt uit $e^{iu(aX+b)} = e^{iauX}e^{iub}$. Neem links en rechts verwachtingen.
(ii) Omdat X en Y onafhankelijk zijn, zijn ook e^{iuX} en e^{iuY} dat. Dus is de verwachting van hun produkt het produkt van hun verwachtingen: $\mathbb{E}e^{iu(X+Y)} = \mathbb{E}e^{iuX}\mathbb{E}e^{iuY}$. □

We geven een paar eenvoudige voorbeelden.

Voorbeeld 1.7. Zij X homogeen verdeeld op $[0,1]$, Zij a en b reële getallen met $a < b$. Dan is $Y = a + (b-a)X$ homogeen verdeeld op $[a,b]$ en dus is $\phi_Y(u) = \frac{e^{iub}-e^{iua}}{i(b-a)u}$. Vergelijk met voorbeeld 1.4 voor de keuze $a = -1$ en $b = 1$.

Karakteristieke functies van andere verdelingen komen in de opgaven aan de orde.

1.2 Een omkeerformule

Er bestaat een algemeen geldende formule om de verdelingsfunctie van een stochastische variabele uit te drukken in zijn karakteristieke functie. We geven hieronder een formule die dat onder zekere voorwaarden doet voor de dichtheidsfunctie.

Veronderstel dat de stochasten X en Y dichtheden f en g hebben (op $\mathbb{R}$) en karakteristieke functies ϕ en ψ. Dan geldt (integralen over $\mathbb{R}$)

$$\int \phi(y)e^{-ity}g(y)\,dy = \int (\int e^{ixy}f(x)\,dx)e^{-ity}g(y)\,dy =$$

$$\int (\int e^{i(x-t)y}g(y)\,dy)f(x)\,dx = \int \psi(x-t)f(x)\,dx.$$

We kiezen nu voor g de dichtheid van de $N(0,a^2)$ verdeling. Dus $g(y) = \frac{1}{a\sqrt{2\pi}}e^{-\frac{1}{2}\frac{y^2}{a^2}}$. Dan is $\psi(t) = e^{-\frac{1}{2}a^2t^2}$ (zie opgave 1.7). Vullen we dit in in bovenstaande vergelijking, dan ontstaat er

$$\int \phi(y)e^{-ity}\frac{1}{a\sqrt{2\pi}}e^{-\frac{1}{2}\frac{y^2}{a^2}}\,dy = \int e^{-\frac{1}{2}a^2(x-t)^2}f(x)\,dx.$$

Vervang nu a in deze vergelijking door $\frac{1}{\sigma}$ en vermenigvuldig daarna met $\frac{1}{\sigma\sqrt{2\pi}}$. Dan komt er

$$\frac{1}{2\pi}\int \phi(y)e^{-ity}e^{-\frac{1}{2}\sigma^2y^2}\,dy = \frac{1}{\sigma\sqrt{2\pi}}\int e^{-\frac{1}{2}\frac{(x-t)^2}{\sigma^2}}f(x)\,dx. \tag{1.2}$$

Vervolgens laten we σ naar nul gaan. Veronderstel dat we limiet en integraal mogen verwisselen, dan convergeert het linkerlid naar

$$\frac{1}{2\pi}\int \phi(y)e^{-ity}\,dy.$$

Het rechterlid van (1.2) stelt de dichtheid voor van $X+Z$, waarbij Z onafhankelijk van X is en $N(0,\sigma^2)$ verdeeld is. (Pas de convolutieformule toe). Als $\sigma \to 0$ dan reduceert dit (Z ontaardt dan in 0) in de dichtheid van X, oftewel f. Dit maakt het volgende plausibel.

Stelling 1.8 *Zij X een stochastische variabele met karakteristieke functie ϕ. Onder de voorwaarde dat $\int |\phi(y)|dy < \infty$, is de functie f gegeven door de omkeerformule*

$$\frac{1}{2\pi}\int \phi(y)e^{-ity}\,dy = f(t)$$

een dichtheid van X. Dus wordt hier de kansverdeling uniek vastgelegd door zijn karakteristieke functie ϕ. Deze dichtheid is bovendien continu.

Het feit dat de kansverdeling uniek wordt vastgelegd door de karakteristieke functies geldt algemener dan in bovenstaande stelling. Wegens het belang hiervan (*karakteristieke functies zijn dus echt karakteristiek*) vermelden we dit als een aparte stelling.

Stelling 1.9 *Laten X en Y twee stochastische variabelen zijn met karakteristieke functies ϕ_X en ϕ_Y. Als geldt dat $\phi_X = \phi_Y$, dan is $X \stackrel{d}{=} Y$.*

Voor discreet verdeelde stochastische variabelen komt een omkeerformule in de opgaven aan bod.

1.3 Opgaven

1.1 Bepaal de karakteristieke functies van de $\text{Bin}(n,p)$ verdeling en van de $\text{Poisson}(\lambda)$ verdeling. Laat zien dat $X+Y$ $\text{Bin}(n+m,p)$ verdeeld is als X en Y onafhankelijk zijn en respectievelijk $\text{Bin}(n,p)$ en $\text{Bin}(m,p)$ verdeeld. Zelfde soort vraag voor twee onafhankelijke Poisson verdeelde stochasten.

1.2 Bepaal de karakteristieke functie van de exponentiële verdeling met parameter λ. Wat is nu de karakteristieke functie van de $\text{Gamma}(n,\lambda)$ verdeling voor gehele n? Opmerking: men kan aantonen dat deze uitdrukking ook klopt voor niet-gehele (positieve) n.

1.3 Bepaal de karakteristieke functie van de dubbelexponentiële verdeling (ook wel Laplace verdeling genaamd), d.w.z. de verdeling met dichtheid $\frac{1}{2}e^{-|x|}$, $x \in \mathbb{R}$.
Laat via de omkeerformule zien dat geldt $\frac{1}{\pi}\int_{-\infty}^{\infty} \frac{1}{1+x^2} e^{itx}\,dx = e^{-|t|}$.

1.4 Bepaal de karakteristieke functie van de (standaard) Cauchy verdeling (gebruik opgave 1.3).
Als $X_1, \ldots, X_n$ onafhankelijk en Cauchy verdeeld zijn, wat is dan de verdeling van $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$. Waarom gaat de zwakke wet van de grote aantallen in dit geval niet op?

1.5 Zij $I(m) = \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{imt}\,dt$, $m \in \mathbb{Z}$. Laat zien dat $I(m) = 1$ als $m = 0$ en $I(m) = 0$ als $m \neq 0$.

1.6 Zij X een stochastische variabele met waarden in $\mathbb{Z}$ en ϕ zijn karakteristieke functie. Laat zien dat

$$P(X = n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \phi(t)e^{-int}\,dt.$$

1.7 De karakteristieke functie van de $N(\mu,\sigma^2)$ verdeling wordt gegeven door $\phi(u) = e^{iu\mu - \frac{1}{2}\sigma^2 u^2}$. Toon dit aan (gebruik voorbeeld 1.5 en propositie 1.6).

2 Stochastische vectoren

Met een stochastische n-vector X bedoelen we (meestal) een kolom-vector bestaande uit stochastische variabelen $X_1, \ldots, X_n$. We kunnen dus voor de getransponeerde schrijven $X^\top = (X_1, \ldots, X_n)$. Echter, bij uitzondering (meestal vanuit typografische overwegingen) zullen we ook wel eens een rij-vector tegenkomen.
Als voor alle X_i geldt dat $\mathbb{E}|X_i| < \infty$, dan definiëren we de verwachtingsvector μ van X als de (kolom-)vector met elementen $\mu_i = \mathbb{E}X_i$.
Op analoge wijze definieert men een stochastische matrix en de verwachting daarvan.

We herinneren ons het begrip covariantie van twee stochastische variabelen. Als X en Y stochasten zijn met $\mathbb{E}X^2$ en $\mathbb{E}Y^2$ eindig, dan is ook $\mathbb{E}|XY|$ eindig omdat $|XY| \leq \frac{1}{2}(X^2 + Y^2)$. Daarom is in deze situatie de covariantie $\mathrm{Cov}(X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)$ goed gedefinieerd en bovendien geldt $\mathrm{Cov}(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y$.
Hiernaast bestaat ook het begrip *correlatiecoefficiënt*. Voor stochastische variabelen X en Y is de correlatiecoefficiënt $\rho(X, Y)$, of kortweg ρ, gedefinieerd als

$$\rho(X, Y) = \frac{\mathrm{Cov}(X, Y)}{\sqrt{\mathrm{Var}(X)\mathrm{Var}(Y)}}. \tag{2.1}$$

Wegens de bekende rekenregels voor variantie en covariantie zien we dat de correlatiecoefficiënt een locatie- en schaalinvariant begrip is: het overgaan van X op $\tilde{X} = aX + b$ (met $a > 0$) geeft dezelfde ρ: $\rho(\tilde{X}, Y) = \rho(X, Y)$.
Deze vaststelling is handig om in te zien dat altijd geldt $-1 \leq \rho \leq 1$. Bij gegeven X en Y noteren we met X' en Y' hun gestandaardiseerde versies, dus bijvoorbeeld is $X' = \frac{X - \mathbb{E}X}{\sqrt{\mathrm{Var}(X)}}$. Dan geldt dus dat $\rho = \mathrm{Cov}(X', Y')$. Verder hebben we de volgende eigenschappen die we als lemma formuleren.

Lemma 2.1 *Er geldt*
(i) $X' - \rho Y'$ *en* Y' *zijn ongecorreleerd.*
(ii) $-1 \leq \rho \leq 1$.

Bewijs. (i) We bekijken $\mathrm{Cov}(X' - \rho Y', Y') = \mathrm{Cov}(X', Y') - \rho \mathrm{Cov}(Y', Y') = \rho - \rho 1 = 0$, omdat zoals opgemerkt wegens standaardisatie $\rho = \mathrm{Cov}(X', Y')$ en $\mathrm{Var}Y' = 1$.
(ii) Omdat voor ongecorreleerde stochastische variabelen de variantie van de som gelijk is aan de som van de varianties, vinden we

$$1 = \mathrm{Var}X' = \mathrm{Var}((X' - \rho Y') + \rho Y') = \mathrm{Var}(X' - \rho Y') + \mathrm{Var}(\rho Y') =$$

$$\mathrm{Var}(X' - \rho Y') + \rho^2 \geq \rho^2.$$

Het gestelde volgt nu. □

We breiden het begrip covariantie nu uit naar het geval we met een stochastische

n-vector X en een stochastische m-vector Y te maken hebben. We veronderstellen $\mathbb{E}X^\top X$ en $\mathbb{E}Y^\top Y$ eindig. Per definitie is nu $\mathrm{Cov}(X,Y)$ de $n \times m$ matrix met als ij-element het reële getal $\mathrm{Cov}(X_i, Y_j)$. We kunnen nu dus schrijven

$$\mathrm{Cov}(X,Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)^\top = \mathbb{E}XY^\top - \mathbb{E}X\mathbb{E}Y^\top. \tag{2.2}$$

We zeggen dat X en Y ongecorreleerd zijn, als hun covariantiematrix nul is. Als geldt dat Y en X hetzelfde zijn, dan schrijven we $\mathrm{Cov}(X)$ i.p.v. $\mathrm{Cov}(X,X)$. Merk op dat $\mathrm{Cov}(X)$ een symmetrisch (vierkante) matrix is.

We geven gelijk een paar eigenschappen. In de onderstaande propositie zijn a en b vectoren en A en B matrices van geschikte afmetingen.

Propositie 2.2 *Laten X en Y stochastische vectoren zijn. Dan geldt*
(i) $\mathbb{E}(AX + a) = A\mathbb{E}X + a$ *en* $\mathrm{Cov}(AX + a, BY + b) = A\mathrm{Cov}(X,Y)B^\top$.
(ii) $\mathrm{Cov}(X)$ *is positief definiet.*

Bewijs. De beweringen onder (i) volgen eenvoudig door uitschrijven.
(ii) We moeten laten zien dat voor elke kolomvector v geldt dat $v^\top \mathrm{Cov}(X)v \geq 0$. Gebruik nu onderdeel (i) met $Y = X$, $A = B = v^\top$ en $a = b = 0$. Dan vinden we $v^\top \mathrm{Cov}(X)v = \mathrm{Cov}(v^\top X)$. Omdat $v^\top X$ reëelwaardig is, is $\mathrm{Cov}(v^\top X)$ niets anders dan de gewone variantie en dus niet-negatief. □

2.1 De transformatieregel voor stochastische vectoren

Vaak zijn we bij een gegeven stochastische variabele of vector X niet alleen geïnteresseerd in zijn eigen verdeling, maar ook in de verdeling van een of andere functie van X, zeg van $Y = g(X)$. In het algemeen is het niet doenlijk om hier een eenvoudige expressie voor af te leiden. Onder beperkende voorwaarden lukt dat wel.
Eerst nog wat terminologie. Als h een vector is van n continu differentieerbare functies h_i van n variabelen, dan bedoelen we met $D_h(y)$ van h in een punt y matrix met als ij-elementen de partiële afgeleiden $\frac{\partial h_i}{\partial y_j}$ uitgerekend in y. $J_h(y)$ (de Jacobiaan van h in y) is dan per definitie de determinant van $D_h(y)$.
Is h de inverse van een functie g, dan gelden de volgende relaties. $D_h(y) = D_g(x)^{-1}$ en $J_h(y) = 1/J_g(x)$ voor $y = g(x)$ (oftewel $x = h(y)$), mits $J_g(x) \neq 0$.

Stelling 2.3 *Zij X een stochastische n-vector met dichtheid f_X. Zij g een op W_X (de waardenverzameling van X) gedefinieerde inverteerbare afbeelding met waarden in $\mathbb{R}^n$ met inverse h. Veronderstel dat g bovendien continu differentieerbaar is en dat de Jacobiaan van g in elk punt van W_X ongelijk nul is. Dan heeft $Y = g(X)$ een dichtheid f_Y gegeven door de formule*

$$f_Y(y) = f_X(h(y))|J_h(y)|. \tag{2.3}$$

Bewijs. Dit is o.m. een gevolg van de substitutieregel voor meerdimensionale integralen. We bewijzen alleen het één-dimensionale geval. Vergelijking (2.3) reduceert nu tot

$$f_Y(y) = f_X(h(y))|h'(y)|. \tag{2.4}$$

Omdat g strikt monotoon is, is h dat ook. Verder is g of stijgend of dalend. We bewijzen (2.4) alleen voor het tweede geval.
Bekijk eerst de verdelingsfunctie van Y.

$$P(Y \leq y) = P(g(X) \leq y) = P(X \geq h(y)) = 1 - F_X(h(y)).$$

De dichtheid f_Y verkrijgen we nu door differentiatie naar y. Doen we dit voor de beide uiterste leden van bovenstaande rij gelijkheden, dan krijgen we m.b.v. de kettingregel

$$f_Y(y) = -f_X(h(y))h'(y) = f_X(h(y))|h'(y)|.$$

omdat $h'(y) \leq 0$ is. □

2.2 Karakteristieke functies voor multivariate verdelingen

In paragraaf 1.1 werden karakteristieke functies gedefinieerd voor eendimensionale verdelingen. In het multivariate geval is dat ook mogelijk, en wel op een wijze die een voor de hand liggende uitbreiding is van wat we eerder gezien hebben.

Definitie 2.4 *Zij X een stochastische n-vector. Dan is de karakteristieke functie ϕ van (de kansverdeling van) X gedefinieerd als de complexwaardige functie op $\mathbb{R}^n$ die geven wordt door $\phi(u) = \mathbb{E}e^{iu^\top X}$. Ook gebruiken we wel de notatie ϕ_X i.p.v. ϕ.*

Een belangrijke rekenregel is (vergelijk propositie 1.2)

$$\phi_{AX+a}(u) = \phi_X(A^\top u)e^{iu^\top a}, \tag{2.5}$$

voor een matrix A en een vector a van geschikte afmetingen. Het bewijs is weer rechttoe rechtaan.

Net als voor stochastische variabelen hebben we ook hier een uniciteitsstelling, die we zonder verdere discussie vermelden.

Stelling 2.5 *Voor stochastische n-vectoren X en Y geldt*

$$X \stackrel{d}{=} Y \Longleftrightarrow \phi_X = \phi_Y. \tag{2.6}$$

Onafhankelijkheid valt nu als volgt te karakteriseren.

Stelling 2.6 *Een stochastische vector heeft onafhankelijke componenten X en Y als en alleen als voor alle u en v geldt $\phi_{(X,Y)}(u,v) = \phi_X(u)\phi_Y(v)$.*

Bewijs. Dit is een eenvoudig gevolg van de uniciteitsstelling. Veronderstel dat X en Y onafhankelijk zijn. Dan valt op dezelfde manier als in propositie 1.2 de factorisatie $\phi_{(X,Y)}(u,v) = \phi_X(u)\phi_Y(v)$ te bewijzen. Omgekeerd, als deze factorisatie geldt dan weten we nu dat deze tot stand komt voor onafhankelijke X en Y. Maar omdat de verdeling van (X,Y) uniek vastgelegd wordt door de karakteristieke functie moeten X en Y dus wel onafhankelijk zijn. □

Met het oog op de volgende paragraaf geven we een belangrijk voorbeeld.

Voorbeeld 2.7. Laten $Z_1, \ldots, Z_n$ *onafhankelijke* stochastische variabelen zijn, alle standaard normaal verdeeld. Zij $X^\top = (Z_1, \ldots, Z_n)$. We bepalen de karakteristieke functie van X. Merk op dat $e^{iu^\top X} = \prod_{k=1}^n e^{iu_k Z_k}$ en dat ook de $e^{iu_k Z_k}$ onafhankelijk zijn. Wederom, net zoals in het bewijs van propositie 1.2, krijgen we dus dat $\mathbb{E}e^{iu^\top X} = \prod_{k=1}^n \mathbb{E}e^{iu_k Z_k} = \prod_{k=1}^n e^{-\frac{1}{2}u_k^2} = e^{-\frac{1}{2}u^\top u}$, waarbij we bij de tweede overgang gebruik hebben gemaakt van de karakteristieke functie van de standaard normale verdeling.
We gaan nu nog een stap verder. Voor een vector a en een matrix A van geschikte afmetingen zetten we $Y = AX + a$. Met behulp van (2.5) vinden we nu $\phi_Y(u) = e^{-\frac{1}{2}u^\top AA^\top u} e^{iu^\top a} = e^{iu^\top a - \frac{1}{2}u^\top AA^\top u}$.

We eindigen deze paragraaf met een lemma.

Lemma 2.8 *Voor elke vector $a \in \mathbb{R}^n$ en elke positief definiete matrix $\Sigma \in \mathbb{R}^{n\times n}$ is er een stochastische n-vector zodanig dat de functie ϕ gegeven door $\phi(u) = e^{iu^\top a - \frac{1}{2}u^\top \Sigma u}$ de karakteristieke functie is van X. Bovendien is de verdeling van X volledig vastgelegd door a en Σ.*

Bewijs. Omdat Σ positief definiet is, kunnen we Σ ontbinden als $\Sigma = U\Lambda U^\top$. Immers, als $\lambda_1, \ldots, \lambda_n$ de niet-negatieve eigenwaarden zijn van Σ en $u_1, \ldots, u_n$ de bijbehorende orthogonale eigenvectoren, dan kiezen we $U = (u_1, \ldots, u_n)$ en Λ als de diagonaal matrix met $\Lambda_{ii} = \lambda_i$.
Zij nu W de diagonaal matrix met $W_{ii} = \sqrt{\lambda_i}$ (hier gebruiken we dus dat de $\lambda_i \geq 0$). Neem nu $A = UW^\top$, dan is $\Sigma = AA^\top$. Pas nu voorbeeld 2.7 toe.
De uniciteit van de verdeling van X volgt uit de uniciteit van karakteristieke functies. $\square$

2.3 Opgaven

2.1 X_1 en X_2 zijn onafhankelijke Gamma(α_1, β) respectievelijk Gamma(α_2, β) verdeelde stochastische variabelen. Bepaal de simultane dichtheid van (Y_1, Y_2) waar $Y_1 = \frac{X_1}{X_1+X_2}$ en $Y_2 = X_1 + X_2$. Wat zijn de marginale verdelingen van de Y_i? Zijn Y_1 en Y_2 onafhankelijk?

2.2 Z is een positieve stochastische variabele met dichtheid $f_Z(z) = ze^{-\frac{1}{2}z^2}$. Onafhankelijk daarvan is Θ uniform verdeeld op $[0, 2\pi]$. Bepaal de simultane dichtheid van (X, Y) met $X = Z\cos\Theta$ en $Y = Z\sin\Theta$.

2.3 Laten X en Y onafhankelijke standaard normaal verdeelde stochastische variabelen zijn. Bepaal met de transformatieregel de simultane dichtheid van (U, V), waar $U = X + Y$ en $V = X - Y$. Zijn U en V onafhankelijk?

2.4 Gegeven zijn twee onafhankelijke stochasten X en Y, beide standaard normaal verdeeld. Zij $g(x, y) = (\frac{x}{y}, y)$. Bepaal de simultane dichtheid van $(U, V) = g(X, Y)$. Wat is de marginale dichtheid van U? Welke oude bekende zien we hierin terug?

2.5 Zij $\alpha, \beta, p > 0$ en bovendien $\alpha > p$. Z en R zijn onafhankelijke stochasten respectievelijk met een Gamma(α, β) en Beta$(p, \alpha - p)$ verdeling. Definieer dan $X = RZ$ en $Y = (1 - R)Z$. Bepaal de simultane dichtheid van X en Y. Zijn X en Y onafhankelijk?

2.6 Veronderstel dat (X, Y) de dichtheid

$$\frac{1}{2\pi\sqrt{1-\rho^2}} e^{-\frac{1}{2(1-\rho^2)}(x^2+y^2-2\rho xy)}$$

heeft. Definieer nieuwe stochasten L $(L \geq 0)$ en Θ $(\Theta \in [0, 2\pi))$ door $X = \sqrt{L}\cos\Theta$ en $Y = \sqrt{L}\sin\Theta$. Bepaal achtereenvolgens de simultane dichtheid van (L, Θ) en de voorwaardelijke dichtheid van L gegeven $\Theta = \theta$. Wat zijn $\mathbb{E}[L|\Theta = \theta]$ en $\mathbb{E}L$?

2.7 Zij X een 3-dimensionale stochastische vector met verwachting $\mu = [2, 1, 3]^\top$ en covariantiematrix $\begin{pmatrix} 4 & 2 & 6 \\ 2 & 1 & 3 \\ 6 & 3 & 9 \end{pmatrix}$. Zij $L = \{v \in \mathbb{R}^3 | \exists \lambda \in \mathbb{R} : v = \lambda\mu\}$. Toon aan dat $P(X \in L) = 1$.

3 De multivariaat normale verdeling

3.1 Definitie en eerste eigenschappen

Een belangrijke reden om in dit college karakteristieke functies in te voeren was om het definiëren van de multivariaat normale verdeling te vergemakkelijken. De definitie is als volgt.

Definitie 3.1 *Een stochastische vector X is multivariaat normaal verdeeld met verwachtingsvector μ en covariantiematrix Σ als de karakteristieke functie van X gegeven wordt door*

$$\phi(u) = e^{iu^\top\mu - \frac{1}{2}u^\top\Sigma u}. \tag{3.1}$$

We gebruiken voor deze verdeling de notatie $N(\mu, \Sigma)$, of $N_n(\mu, \Sigma)$ als we willen benadrukken dat X een n-vector is. In het geval $n = 2$ spreken we ook wel van de bivariaat normale verdeling.

Dat de multivariaat normale verdeling door deze definitie eenduidig wordt vastgelegd volgt onmiddellijk uit lemma 2.8. Direct uit de definitie volgt nu een belangrijke karakterisering die we als stelling formuleren.

Stelling 3.2 *Een stochastische vector X heeft een $N_n(\mu, \Sigma)$ verdeling als en alleen als alle lineaire combinaties van deze vector normaal verdeeld zijn. Specifiek: als X een $N_n(\mu, \Sigma)$ verdeling heeft, dan heeft voor elke $a \in \mathbb{R}^n$ de stochastische variabele $a^\top X$ een $N(a^\top\mu, a^\top\Sigma a)$ verdeling en vice versa.*

Bewijs. Veronderstel dat X een $N_n(\mu, \Sigma)$ verdeling heeft en zij $Y = a^\top X$. Dan is $\phi_Y(u) = \phi_X(ua) = e^{iua^\top \mu - \frac{1}{2}u^2 a^\top \Sigma u}$, de karakteristieke functie van de $N(a^\top \mu, a^\top \Sigma a)$ verdeling. Wegens de uniciteitsstelling is dus Y zodanig verdeeld.
Veronderstel nu omgekeerd dat alle lineaire combinaties variabelen $a^\top X$ een $N(a^\top \mu, a^\top \Sigma a)$ verdeling hebben. We bepalen nu de karakteristieke functie van X in het punt a: $\phi_X(a) = \mathbb{E}e^{ia^\top X} = \phi_{a^\top X}(1) = e^{ia^\top \mu - \frac{1}{2}a^\top \Sigma a}$, juist de karakteristieke functie van de $N_n(\mu, \Sigma)$ verdeling, zodat X wegens de uniciteitsstelling dus inderdaad deze verdeling heeft. $\square$

De karakterisering van stelling 3.2 stelt ons in staat om een aantal gevolgen eenvoudig aan te geven. Zo is bijvoorbeeld elke component X_i van een multivariaat verdeelde X normaal verdeeld, omdat $X_i = e_i^\top X$, met e_i de i-de basisvector. Let op: het is **niet** noodzakelijk zo, dat een stochastische vector X multivariaat normaal verdeeld, als alle componenten afzonderlijk normaal verdeeld zijn! Zie hiervoor opgave 3.2.
Wel is het volgende waar: als $X_1, \ldots, X_n$ *onafhankelijke* stochastische variabelen zijn, dan weten we dat alle lineaire combinaties $\sum a_i X_i$ weer normaal verdeeld zijn. In dit geval is de vector X met deze X_i als elementen dus multivariaat normaal verdeeld.
Ook elke deelvector van een multivariaat normaal verdeelde stochastische vector X is weer multivariaat verdeeld. Immers, elke deelvector Y van X kunnen we schrijven als $Y = HX$ voor een geschikte matrix H bestaande uit enen en nullen. (Hierboven hebben we al het speciale geval gezien dat $X_i = HX$ met $H = e_i^\top$). Een lineaire combinatie van de elementen van Y kunnen we schrijven als $b^\top Y$. Maar dit is gelijk aan $a^\top X$ met $a^\top = b^\top H$ en dus ook een lineaire combinatie van de elementen van X, en daarmee normaal verdeeld. Deze eigenschap kan als volgt worden uitgebreid.

Propositie 3.3 *Als X een multivariate $N(\mu, \Sigma)$ verdeling heeft, a een vector is en A een matrix, dan is $Y = AX + a$ multivariaat $N(A\mu + a, A\Sigma A^\top)$ verdeeld.*

Bewijs. Uit propositie 2.2 volgt dat $\mathbb{E}Y = A\mu + a$ en dat $\mathrm{Cov}(Y) = A\Sigma A^\top$. We laten nu zien dat Y multivariaat normaal verdeeld is. Neem een lineaire combinatie $c^\top Y$. Schrijf dit als $(c^\top A)X + c^\top a$. Nu is $(c^\top A)X$ een lineaire combinatie van de elementen van X en dus normaal verdeeld. Tellen we daar het reële getal $c^\top a$ bij op, dan hebben we nog steeds een normaal verdeelde stochast. Elke lineaire combinatie $c^\top Y$ is dus normaal verdeeld, waaruit het gestelde volgt wegens stelling 3.2. $\square$

Het is eenvoudig in te zien dat een stochastische vector met covariantiematrix Σ geen dichtheid kan hebben als Σ singulier is. We kunnen ons wel afvragen wat de dichtheid van een $N_n(\mu, \Sigma)$ verdeelde stochastische vector is, als Σ niet-singulier is. Het antwoord staat hieronder.

Propositie 3.4 *Veronderstel dat Σ een niet-singuliere positief definiete matrix* *is. Een stochastische vector X heeft een $N_n(\mu, \Sigma)$ verdeling als en alleen als X de dichtheid*

$$f_X(x) = \frac{1}{(2\pi)^{n/2}(\det \Sigma)^{1/2}} e^{-\frac{1}{2}(x-\mu)^\top \Sigma^{-1}(x-\mu)} \tag{3.2}$$

heeft.

Bewijs. Schrijf net als in het bewijs van lemma 2.8 $\Sigma = AA^\top$. In dit geval is A inverteerbaar. Zij nu Z een vector van onafhankelijke standaard normale stochastische variabelen. Dan is Z multivariaat normaal verdeeld en wegens propositie 3.3 is $AZ + \mu$ juist $N(\mu, \Sigma)$ verdeeld is. Het is dus voldoende om de dichtheid van $AZ + \mu$ te bepalen.
Merk nu op dat de transformatie $g(z) = Az + \mu$ inverteerbaar is met inverse $h(x) = A^{-1}(x-\mu)$. We passen nu de transformatiestelling toe. Daarvoor hebben we de matrix van partiële afgeleides van h nodig. Het is makkelijk in te zien dat $\frac{\partial h_i}{\partial x_j}$ precies gelijk is aan het ij-element van A^{-1}. Verder is $\det A^{-1} = \frac{1}{\det A}$ en $\det A = (\det \Sigma)^{1/2}$. Omdat Z de dichtheid $f_Z(z) = \frac{1}{(2\pi)^{n/2}} e^{-\frac{1}{2}z^\top z}$ heeft, vinden we door toepassing van (2.3) dat $AZ + \mu$ de dichtheid van vergelijking (3.2) heeft, omdat $(A^{-1})^\top A^{-1} = (AA^\top)^{-1} = \Sigma^{-1}$.

3.2 Onafhankelijkheid

Het is bekend dat twee stochastische variabelen ongecorreleerd zijn, zo gauw ze onafhankelijk zijn (mits de varianties bestaan), maar ook dat het omgekeerde i.h.a. *niet waar* is. Het bijzondere van multivariaat normale verdelingen is dat een omkering wel van kracht is. Ter voorbereiding van een precieze formulering eerst een beetje boekhouden.

Veronderstel dat we te maken hebben met twee stochastische vectoren X_1 (n-dimensionaal) en X_2 (m-dimensionaal). Schrijf X voor de gezamenlijke vector, dus $X^\top = (X_1^\top, X_2^\top)$. Veronderstel verder dat alle covarianties (en dus verwachtingen bestaan). Schrijf $\mu_1 = \mathbb{E}X_1$, $\mu_2 = \mathbb{E}X_2$, $\mu = \mathbb{E}X$ en verder $\Sigma_{11} = \mathrm{Cov}(X_1)$, $\Sigma_{22} = \mathrm{Cov}(X_2)$, $\Sigma_{12} = \mathrm{Cov}(X_1, X_2)$, $\Sigma_{21} = \mathrm{Cov}(X_2, X_1)$ en $\sigma = \mathrm{Cov}(X)$. Dan geldt dat $\Sigma_{12} = \Sigma_{21}^\top$ en

$$\mu = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Merk nog even op dat X_1 en X_2 ongecorreleerd zijn als $\Sigma_{12} = 0$. In dat geval is de matrix Σ dus blokdiagonaal.

Stelling 3.5 *Veronderstel dat X een $N(\mu, \Sigma)$ verdeling heeft. Dan hebben X_1 en X_2 een $N(\mu_1, \Sigma_{11})$ respectievelijk $N(\mu_2, \Sigma_{22})$ verdeling. Bovendien geldt dat X_1 en X_2 onafhankelijk zijn als en alleen als ze ongecorreleerd zijn.*

Bewijs. De normaliteit van X_1 en X_2 zijn we al tegengekomen in paragraaf 3.1. Ook weten we al dat onafhankelijkheid ongecorreleerdheid impliceert. Het enige dat nog aangetoond moet worden is dat ongecorreleerd van X_1 en X_2 in *deze situatie* hun onafhankelijkheid impliceert.
Welnu, de karakteristieke functie van X is zoals we weten (definitie 3.1) gegeven door $\phi_X(u) = e^{iu^\top \mu - \frac{1}{2}u^\top \Sigma u}$ en die van elk van de X_k is (wegens dezelfde definitie) gegeven door $\phi_{X_k}(u_k) = e^{iu_k^\top \mu_k - \frac{1}{2}u_k^\top \Sigma_{kk} u_k}$. Met $u^\top = (u_1^\top, u_2^\top)$ zien we nu dat $\phi_X(u) = \phi_{X_1}(u_1)\phi_{X_2}(u_2)$ waarbij we gebruik maken van de blokdiagonaliteit van Σ. Het gestelde volgt nu door toepassing van stelling 2.6. □

Opmerking 3.6. Stelling 3.5 laat zich eenvoudig uitbreiden tot een splitsing van X in meer dan twee componenten. In het bijzonder indien Σ een echte diagonaalmatrix is, dan zijn *alle* componenten van X onderling onafhankelijk.

Opmerking 3.7. We benadrukken nog eens apart dat het in stelling 3.5 van essentieel belang is dat de *hele* vector X een multivariaat normale verdeling heeft.

3.3 Conditioneren bij een multivariaat normale verdeling

Zij X een multivariaat normaal verdeelde stochastische vector. Schrijf $X = \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, en verder $\mathbb{E}X = \begin{pmatrix} \mathbb{E}X_1 \\ \mathbb{E}X_2 \end{pmatrix} = \begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}$ en $\mathrm{Cov}(X) = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}$.
Voor het gemak nemen we in deze paragraaf aan dat $\Sigma = \mathrm{Cov}(X)$ *niet-singulier* is. Dan zijn ook Σ_{11} en Σ_{22} niet-singulier. Het is mogelijk om ook zonder deze aanname te werken, de prijs die hiervoor betaald moet worden is wat ingewikkeldere Lineaire Algebra.

Eerst een technisch hulpmiddel. Hoewel eenvoudig te bewijzen is de decompositie in onderstaand lemma vaak handig om iets uit te rekenen.

Lemma 3.8 *Zij $A \in \mathbb{R}^{m\times m}, B \in \mathbb{R}^{n\times m}, C \in \mathbb{R}^{m\times n}, D \in \mathbb{R}^{n\times n}$ en bovendien D inverteerbaar. Dan geldt*

$$\begin{pmatrix} A & C \\ B & D \end{pmatrix} = \begin{pmatrix} I & CD^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} A - CD^{-1}B & 0 \\ 0 & D \end{pmatrix} \begin{pmatrix} I & 0 \\ D^{-1}B & I \end{pmatrix}. \tag{3.3}$$

Bewijs. Uitschrijven. □

Lemma 3.9 *Zij $Y = X_1 - \Sigma_{12}\Sigma_{22}^{-1}X_2$.*
Dan is $\begin{pmatrix} Y \\ X_2 \end{pmatrix}$ multivariaat normaal met verwachting $\begin{pmatrix} \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}\mu_2 \\ \mu_2 \end{pmatrix}$ en covariantiematrix $\begin{pmatrix} \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21} & 0 \\ 0 & \Sigma_{22} \end{pmatrix}$.
Bovendien zijn Y en X_2 onafhankelijk.

Bewijs. Merk op dat $\begin{pmatrix} Y \\ X_2 \end{pmatrix} = \begin{pmatrix} I & -\Sigma_{12}\Sigma_{22}^{-1} \\ 0 & I \end{pmatrix} \begin{pmatrix} X_1 \\ X_2 \end{pmatrix}$, een lineaire transformatie is van X, en dus weer multivariaat normaal verdeeld is. De covariantiematrix volgt eenvoudig.
In het bijzonder is de matrix $\mathrm{Cov}(Y, X_2)$ nul. Derhalve zijn Y en X_2 niet gecorreleerd en daarmee ook onafhankelijk. □

Stelling 3.10 *Zij $x_2 \in \mathbb{R}$. Dan is de voorwaardelijke verdeling van X_1 gegeven $X_2 = x_2$ (multivariaat) normaal met verwachting $\mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$ en covariantiematrix $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.*

Bewijs. Wegens het voorgaande lemma zijn Y en X_2 onafhankelijk en dus zijn ook $\tilde{Y} = Y + \Sigma_{12}\Sigma_{22}^{-1}x_2$ en X_2 dat.
Maar dan geldt voor willekeurige B wegens de genoemde onafhankelijkheid

$$P(\tilde{Y} \in B) = P(\tilde{Y} \in B | X_2 = x_2) =$$

$$P(X_1 - \Sigma_{12}\Sigma_{22}^{-1}(X_2 - x_2) \in B | X_2 = x_2) = P(X_1 \in B | X_2 = x_2),$$

waarbij de laatste gelijkheid geldt omdat (triviaal) $X_2 - x_2 = 0$ onder de voorwaarde $X_2 = x_2$.
Merk nu op dat $\tilde{Y}$ multivariaat normaal verdeeld is met $\mathbb{E}\tilde{Y} = \mu_1 - \Sigma_{12}\Sigma_{22}^{-1}(\mu_2 - x_2)$ en $\mathrm{Cov}(\tilde{Y}) = \mathrm{Cov}(Y)$, waarmee het bewijs rond is. □

Opmerking 3.11. We zien in bovenstaande stelling dat de voorwaardelijke verwachting $\mathbb{E}[X_1 | X_2 = x_2] = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$, een grootheid afhankelijk van x_2. Daarentegen hangt de covariantiematrix van de voorwaardelijke verdeling *niet* van x_2 af.

3.4 Opgaven

3.1 Bewijs de factorisatie in lemma 3.8.

3.2 Zij (X, Y) een twee-dimensionale stochastische vector met kansdichtheid

$$f(x, y) = \begin{cases} \frac{1}{\pi} e^{-\frac{1}{2}(x^2+y^2)} \text{ als } xy > 0 \\ 0 \text{ anders} \end{cases}$$

Laat zien dat zowel X als Y standaard normaal verdeeld zijn. Is (X, Y) multivariaat (twee-dimensionaal) normaal verdeeld?

3.3 Laat zien dat de uitdrukking voor de covariantiematrix in lemma 3.9 correct is.

3.4 Zij X multivariaat normaal verdeeld met niet-singuliere covariantie matrix Σ. Schrijf $X^\top = [X_1^\top, X_2^\top]$, etc. zoals in het begin van deze paragraaf. Veronderstel dat X_1 waarden in $\mathbb{R}^m$ aanneemt. Bepaal via stelling 3.10 de voorwaardelijke dichtheid van X_1 gegeven $X_2 = x_2$: Met $\hat{\mu}_1 = \mu_1 + \Sigma_{12}\Sigma_{22}^{-1}(x_2 - \mu_2)$ en $\hat{\Sigma}_{11} = \Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$ geldt

$$f_{X_1|X_2=x_2}(x_1) = (2\pi)^{-m/2}(\det \hat{\Sigma}_{11})^{-\frac{1}{2}} \exp(-\frac{1}{2}(x_1 - \hat{\mu}_1)^\top \hat{\Sigma}_{11}^{-1}(x_1 - \hat{\mu}_1)).$$

3.5 Zelfde situatie als in opgave 3.4. Bepaal nu via de formule van voorwaardelijke dichtheden de voorwaardelijke dichtheid van X_1 gegeven $X_2 = x_2$.
Aanwijzing: gebruik lemma 3.8 om eerst Σ en dan Σ^{-1} te schrijven als produkt van drie factoren.

3.6 Onderstel dat X en Y de volgende simultane dichtheid op $\mathbb{R}^2$ hebben.

$$f(x,y) = c \cdot e^{-(1+x^2)(1+y^2)},$$

waarbij c natuurlijk zo gekozen is dat de integraal van f gelijk is aan 1.
Laat zien dat $c\sqrt{\frac{\pi}{1+x^2}}e^{-(1+x^2)}$ de marginale dichtheid van X voorstelt.
Toon vervolgens aan dat Y gegeven $X = x$ *normaal* verdeeld is met variantie $\frac{1}{2(1+x^2)}$. Hoe kun je hieruit concluderen dat (X, Y) niet bivariaat normaal verdeeld is?

3.7 Laat (X, Y) bivariaat normaal verdeeld zijn. Veronderstel dat de variantie van Y niet nul is. Zij $f(a, b) = \mathbb{E}(X - aY - b)^2$. Voor welke a en b heeft f een minimum? Wat stelt $aY + b$ voor deze waarden van a en b voor?

3.8 Onderstel dat de stochastische 2-vector X de karakteristieke functie gegeven door $\phi(u_1, u_2) = \exp(iu_1 + 2iu_2 - \frac{1}{2}u_1^2 + 2u_1u_2 - 6u_2^2)$ heeft. Bepaal dan de verdeling van X.

3.9 Veronderstel dat (X, Y) bivariaat normaal verdeeld, zo dat X en Y elk standaard normaal zijn en correlatiecoëfficient $\rho \in (-1, +1)$ hebben.
Laat zien dat X en $Y - \rho X$ onafhankelijk zijn. (Vergelijk met lemma 2.1).

3.10 Veronderstel dat (X, Y) bivariaat normaal verdeeld, zo dat X en Y gelijke variantie hebben. Laat zien dat $X + Y$ en $X - Y$ onafhankelijk zijn.

3.11 Veronderstel dat (X, Y) bivariaat normaal verdeeld, zo dat X en Y elk standaard normaal zijn en correlatiecoëfficient $\rho \in (-1, +1)$ hebben.
Bepaal $\mathbb{E}[Y|X = x]$ en $\mathrm{Var}(Y|X = x)$.

3.12 Zij X multivariaat normaal verdeeld met $\mathbb{E}X = 0$ en covariantiematrix

$$\begin{pmatrix} 1 & 2 & -1 \\ 2 & 6 & 0 \\ -1 & 0 & 4 \end{pmatrix}.$$

Definieer $Y = AX$ met

$$A = \begin{pmatrix} 1 & 0 & -1 \\ 2 & -1 & 0 \\ 0 & -1 & 2 \end{pmatrix}.$$

Bepaal de voorwaardelijke verdeling van Y_3 gegeven $Y_1 = 0$ en $Y_2 = 1$.