# 5
# Continue informatie

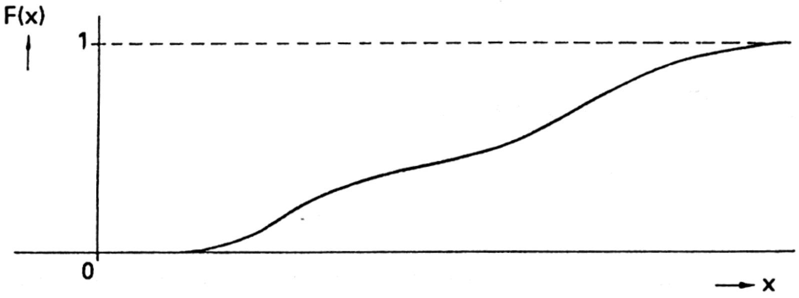
## 5.1. Kansdichtheden

Alvorens informatiematen voor het continue geval te introduceren dient allereerst enige aandacht besteed te worden aan het begrip *kansdichtheid.* In het discrete geval kan een stochastische variabele verkregen worden door de uitkomsten van een experiment, gerepresenteerd door punten in een uitkomstenruimte, af te beelden op een getallenrechte en vervolgens aan elk getal op deze getallenrechte een kans te verbinden. Bij een continue stochastische variabele is er sprake van oneindig veel uitkomsten in een (eindige of oneindige) uitkomstenruimte. Op deze uitkomstenruimte zijn weer gebeurtenissen te definiëren, waaraan een bepaalde kans verbonden kan worden. Door de uitkomsten van de uitkomstenruimte af te beelden op de getallenrechte via een gegeven functie ontstaat een continue stochastische variabele, waarvan het waardenbereik een (eindig of oneindig) interval op de getallenrechte is.
Het vinden van de kansen op een bepaalde uitkomst, dus op een bepaalde waarde van de continue stochastische variablele is niet zo eenvoudig als in het geval van een discrete stochastische variabele. Omdat er namelijk oneindig veel waarden zijn, moet de kans op elke waarde nul zijn, wil men kunnen verwachten dat de kans op de zekere gebeurtenis één is. Wel kan men echter spreken over de kans dat een waarde ligt in een bepaald deelinterval van de getallenrechte. Dit laatste komt overeen met het begrip gebeurtenis. Zo zal men van een spanning die bijvoorbeeld ligt tussen –10 V en +10 V kunnen zeggen dat de kans op exact 6 V nul is, terwijl de kans dat de spanning ligt tussen 5,9 V en 6,12 V heel goed 0,8 zou kunnen zijn en dus > 0 is.
Gebeurtenissen van de vorm ($a < \mathbf{x} \leq b$) en dergelijke zijn goed te bepalen via een distributiefunctie. En distributiefunctie geeft de kans weer dat een stochastische variabele een waarde tot aan een bepaalde drempelwaarde aanneemt:

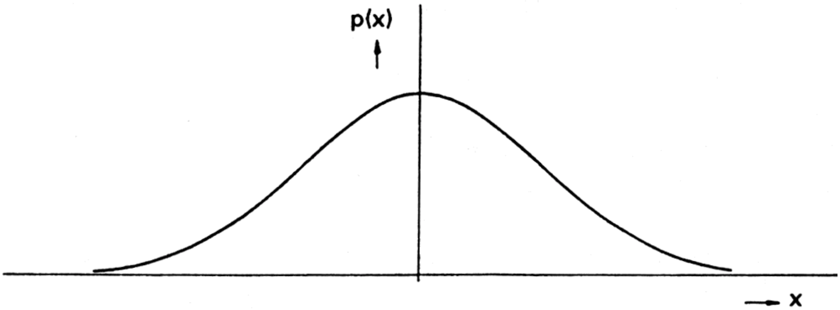
$$F(x) = P(\mathbf{x} \leq x).$$

Voor continue stochastische variabelen is de distributiefunctie een continue functie. In figuur 5.1 is een voorbeeld gegeven van een distributiefunctie:
Als de distributiefunctie een afgeleide bezit in het punt $x$, noemen we deze afgeleide de *kansdichtheid* $p(x)$:

$$p(x) = \frac{\mathrm{d}\,F(x)}{\mathrm{d}x},$$




*Figuur 5.1. Voorbeeld van een distributiefunctie.*



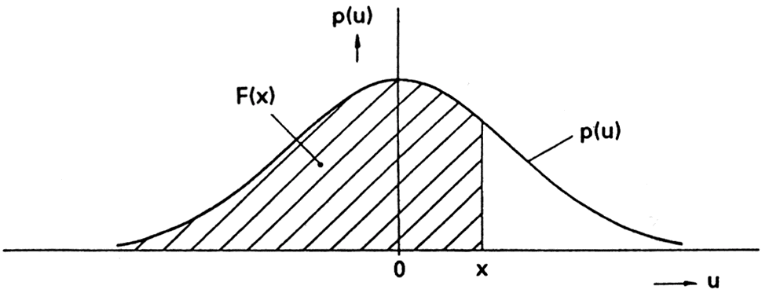

*Figuur 5.2.Voorbeeld van een kansdichtheid.*

waarbij dus moet gelden dat $F(x)$ differentieerbaar is. In figuur 5.2 is een voorbeeld gegeven van een kansdichtheid:
Van de kansdichtheid zullen enkele eigenschappen worden gegeven. Uit het feit dat de distributiefunctie monotoon niet-dalend is volgt direct de eigenschap dat:

$$p(x) \geq 0.$$

Verder geldt:

$$\int_{-\infty}^{x} p(u)\,\mathrm{d}u = F(x) - F(-\infty) = F(x),$$

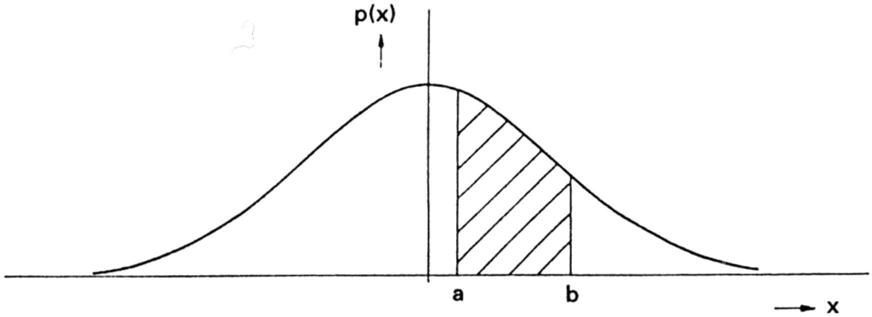



*Figuur 5.3.*

waarbij de definitie van $p(x)$ is gebruikt. Hieruit blijkt dat de waarde van de distributiefunctie overeenkomt met het in figuur 5.3 gearceerde oppervlak:
Hieruit is direct te zien dat geldt:

$$\int_{-\infty}^{\infty} p(x)\,dx = F(\infty) - F(-\infty) = 1. \tag{5.1}$$

Het oppervlak onder de gehele kromme moet dus gelijk zijn aan één. Neemt men als intergratiegrenzen a en b, dan geldt (zie figuur 5.4):




*Figuur 5.4.*

$$\int_{a}^{b} p(x)\,dx = F(b) - F(a) = P(\underline{x} \leq b) - P(\underline{x} \leq a) = P(a < \underline{x} \leq b).$$

De kans dat een continue stochastische variabele een waarde aanneemt tussen $a$ en $b$ is dus te vinden door zijn kansdichtheid te integreren over $(a,b)$. Als $a = b$ volgt dat $P(\mathbf{x} = a) = 0$ hetgeen overeenstemt met de reeds eerder gemaakte opmerking. De kansdichtheid van een stochastische variabele vervult een rol die zeer veel lijkt op de rol die de kansen vervullen bij discretre stochastische variabelen. Bij een continue stochastische variabele is er slechts dan sprake van een kans indien de kansdichtheid wordt geïntegreerd over een bepaald interval $(a,b)$. De kansdichtheid kan in verband daarmee ook worden opgevat als een limiet. Voor een kleine waarde van $\Delta x$ volgt namelijk:

$$P(x < \mathbf{x} \leq x + \Delta x) \approx p(x)\,\Delta x,$$

omdat het oppervlak onder de kansdichtheidskromme over het interval $(x, x + \Delta x)$ te benaderen is met een rechthoek ter breedte $\Delta x$. Hieruit volgt:

$$p(x) = \lim_{\Delta x \to 0} \frac{P(x < \underline{x} \leq x + \Delta x)}{\Delta x}.$$

Een kansdichtheid kan groter dan één zijn, hetgeen eveneens een verschil vormt met een kans, waarvoor immers geldt dat deze ten hoogste één kan zijn. Wel geldt dat de integraal van $p(x)$ over een bepaald interval, dus de kans dat **x** een waarde x aanneemt in dit interval, kleiner dan of gelijk aan één is. Een bekende continue kansverdeling is de uniforme verdeling. Een continue stochastische variabele **x** bezit een uniforme verdeling, indien voor de kansdichtheid geldt:

$$p(x) = \frac{1}{b-a} \qquad a \le x \le b; \tag{5.2}$$

$$= 0 \qquad x < a,\ x > \mathrm{b}.$$

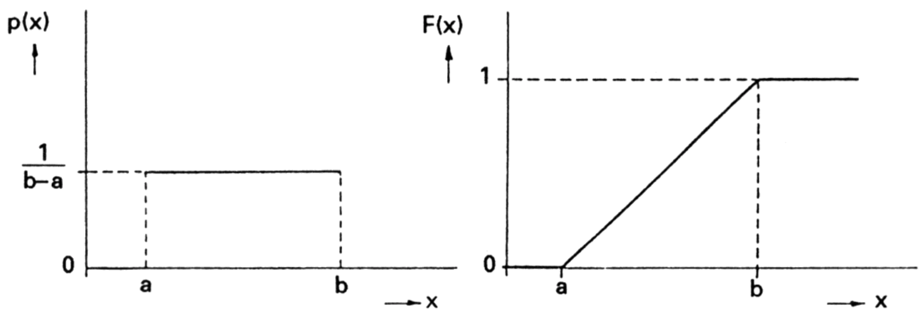
De distributiefunctie van **x** is:

$$F(x) = 0 \qquad x < a;$$

$$= \frac{x-a}{b-a} \qquad a \le x \le b;$$

$$= 1 \qquad x > b.$$

Merk op dat voor $(b - a) < 1$ de kansdichtheid groter is dan één is. De functies $p(x)$ en $F(x)$ zijn weergegeven in figuur 5.5:




*Figuur 5.5. Kansdichtheidsfunctie en distributiefunctie bij een uniforme verdeling.*

Een andere bekende continue kansverdeling is *de normale of gaussische verdeling*. Deze verdeling is belangrijk omdat men van veel verschijnselen met goede benadering kan zeggen dat er sprake is van normaal verdeelde stochastische variabelen, zoals bij ruis in communicatiesystemen en bij meetfouten die ontstaan bij waarnemingen aan systemen en signalen. Een stochastische variabele **x** bezit een normale of gaussische verdeling indien de kansdichtheid wordt gegeven door:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\} \qquad -\infty < x < \infty. \tag{5.3}$$

Hierin zijn $\mu$ en $\sigma$ twee parameters. Met betrekking tot continue verdelingen wordt vaak gebruik gemaakt van het *gemiddelde of de verwachting* van **x**, gedefinieerd door:

$$E(x) = \int_{-\infty}^{\infty} x\, p(x)\, \mathrm{d}x,$$

en van de *variantie*. De variantie is een maat voor de variaties van de mogelijke waarden van $x$ om zijn gemiddelde, en wordt gedefinieerd door:

$$\mathrm{var}(x) = E\left[(x - E(x))^2\right] = \int_{-\infty}^{\infty} (x - E(x))^2 \cdot p(x)\, \mathrm{d}x.$$

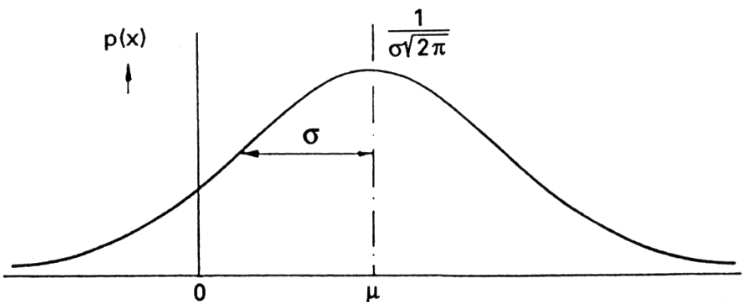
Berekenen we nu $E(x)$ en $\mathrm{var}(x)$ voor de gaussische verdeling dan vinden we

$$E(x) = \mu$$

en

$$\mathrm{var}(x) = \sigma^2.$$

Hieruit volgt dus dat de parameter $\mu$ juist de verwachting van $x$ voorstelt en $\sigma^2$ de variantie. Volledigheidshalve zij vermeld dat de vierkantswortel uit de variantie de *standaardafwijking, standaarddeviatie of spreiding* wordt genoemd en in dit geval dus gelijk is aan $\sigma$. De parameters $\mu$ en $\sigma$ zijn kenmerkend voor de vorm van $p(x)$. Vaak wordt de verdeling dan ook wel aangeduid met $N(\mu, \sigma)$. De parameter $\mu$ geeft het symmetriepunt van de grafiek van $p(x)$ aan en $\mu - \sigma$ en $\mu + \sigma$ de buigpunten van deze grafiek. (Zie figuur 5.6).

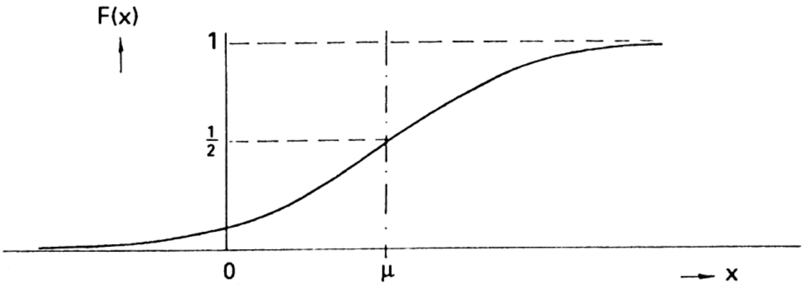



*Figuur 5.6. Gaussische verdeling.*

De hierbij behorende distributiefunctie (zie figuur 5.7) is:

$$F(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left\{-\frac{(w-\mu)^2}{2\sigma^2}\right\} \mathrm{d}w.$$

Voor een gaussiche distributiefunctie $F(x)$ met $\mu = 0$ en $\sigma = 1$ bestaan tabellen. Een willekeurige gaussische verdeling $N(\mu, \sigma)$ is te normeren tot een verdeling $N(0,1)$ door de transformatie $\mathbf{y} = (\mathbf{x} - \mu)/\sigma$ toe te passen op de stochastische variabele $\mathbf{x}$. De verdeling $N(0,1)$ wordt gewoonlijk standaardnormaal genoemd. Uit de symmetrie van de kansdichtheid volgt dat $p(\mu + a) = p(\mu - a)$ en tevens dat $F(\mu) = 1/2$, omdat dan juist over het halve waardenbereik van **x** geïntegreerd is.




*Figuur 5.7. Distributie van gaussische verdeling.*

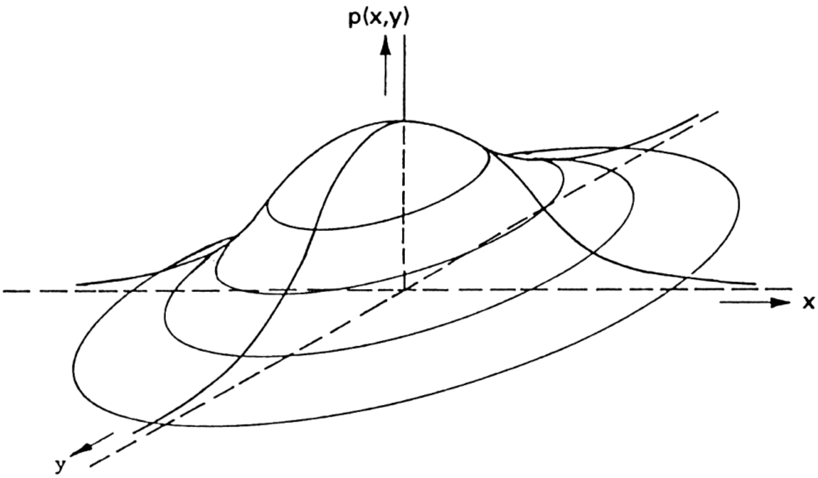
Naast het beschouwen van één continue stochastische variabele zal men ook te maken hebben met combinatie van continue stochastische variabelen. Stel dat er sprake is van twee continue stochastische variabelen **x** en **y** met kansdichtheden $p(x)$ en $p(y)$. Voor elk getalpaar $(x,y)$ zijn we nu geïnteresseerd in $p(x,y)$. Om hiertoe te komen wordt allereerst de gezamenlijke distributiefunctie beschouwd voor twee continue stochastische variabelen **x** en **y.** Deze is gedefinieerd als:

$$F(x,y) = P(\mathbf{x} \leq x, \mathbf{y} \leq y).$$

De *gezamenlijke* of *twee-dimensionale* kansdichtheid $p(x,y)$ wordt gedefinieerd als de partiële afgeleide naar $x$ en $y$:

$$p(x,y) = \frac{\partial^2 F(x,y)}{\partial \mathrm{x}\, \partial y} .$$

In figuur 5.8 is een voorbeeld gegeven van een twee-dimensionale kansdichtheid:




*Figuur 5.8. Voorbeeld van twee-dimensionale kansdichtheid.*

Ook nu zal gelden dat:

$$\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1,$$

hetgeen inhoudt dat het volume onder de gezamenlijke kansdichtheid één moet zijn. En juist zoals er in het discrete geval een relatie was aan te geven tussen de marginale en gezamenlijke kans geldt nu

$$p(x) = \int_{-\infty}^{\infty} p(x,y)\,\mathrm{d}y,$$

$$q(y) = \int_{-\infty}^{\infty} p(x,y)\,\mathrm{d}x.$$

Als voorbeeld van een twee-dimensionale kansdichtheid beschouw de twee-dimensionale gaussische verdeling. De uitdrukking hiervoor is:

$$p(x_1,x_2) = \frac{1}{2\pi\sigma_1\sigma_2\sqrt{1-\rho^2}} \cdot \exp\left\{-\frac{1}{2(1-\rho^2)}\left[\frac{(x_1-\mu_1)^2}{\sigma_1^2} + {} - \frac{2\rho(x_1-\mu_1)\cdot(x_2-\mu_2)}{\sigma_1\sigma_2} + \frac{(x_2-\mu_2)^2}{\sigma_2^2}\right]\right\}, \tag{5.4}$$

met parameters $\mu_1$ , $\mu_2$, $\sigma_1$, $\sigma_2$ en $\rho$. Hierbij is $\rho$ de correlatiecoëfficient, waarvoor geldt $-1 < \rho < 1$ en waarover straks meer. De kansdichtheid is weergegeven in figuur 5.8 voor het geval $\mu_1 = \mu_2 = 0$, $\sigma_1 = \sigma_2$, $\rho = 0$. De marginale kansdichtheden $p(x)$ en $p(y)$ zijn ook normaal verdeeld volgens $N(\mu_1, \sigma_1)$, resp. $N(\mu_2,\sigma_2)$.
Als er bij continue stochastische variabelen sprake is van een onderlinge afhankelijkheid zal men gebruik kunnen maken van *conditionele kansdichtheden*. Het gebruik van conditionele kansdichtheden is van groot belang bij problemen uit de elektro- en informatietechniek. Dit komt omdat ruissignalen een continue waardenbreik bezitten, waardoor problemen zoals het scheiden van informatiedragende signalen en ruissignalen en het interpreteren van door ruis gestoorde metingen van nature worden geformuleerd als een conditionele kansdichtheid (of kans), waarbij de conditie wordt gevormd door de gemeten uitkomsten. Voor de definitie van de conditionele kansdichtheid kan, in analogie met het discrete geval, uitgegaan worden van de marginale en de gezamenlijke kansdichtheid. Er geldt:

$$p(x/y) = \frac{p(x,y)}{q(y)}\ .$$

Het verband tussen de marginale kansdichtheid $p(x)$ en de conditionele kansdichtheid

$p(x/y)$ wordt gegeven door:

$$p(x) = \int_{-\infty}^{\infty} p(x,y)\,\mathrm{d}y = \int_{-\infty}^{\infty} q(y)\cdot p(x/y)\,\mathrm{d}y.$$

De formule van Bayes kan eveneens worden gegeven voor kansdichtheden en luidt dan:

$$p(x/y) = \frac{p(x)\cdot p(y/x)}{q(y)}\,, \tag{5.5}$$

of ook:

$$p(x/y) = \frac{\mathrm{p}(x)\cdot p(y/x)}{\int_{-\infty}^{\infty} p(x)\cdot p(y/x)\,\mathrm{d}x}\,.$$

Volledigheidshalve zij vermeld dat de laatste drie uitdrukkingen ook gelden indien **x** een continue en **y** een discrete stochastische variabele is, waarbij deze uitdrukkingen dan de volgende gedaante krijgen:

$$p(x) = \sum_{\text{alle } y_\mathrm{i}} P(y_\mathrm{i})\cdot p(x/y_\mathrm{i}),$$

$$p(x/y_\mathrm{i}) = \frac{p(x)\cdot \mathrm{Q}(y_i/x)}{\mathrm{Q}(y_\mathrm{i})}\,,$$

$$\mathrm{p}(x/y_i) = \frac{p(x)\cdot \mathrm{Q}(y_i/x)}{\int_{-\infty}^{\infty} p(x)\cdot \mathrm{Q}(y_i/x)\,\mathrm{d}x}\,.$$

In de informatietheorie en in het bijzonder in de statistische detectietheorie wordt van deze relaties gebruik gemaakt. In analogie met het discrete geval kan stochastische onafhankelijkheid van twee continue stochastische variabelen **x** en **y** worden gedefinieerd als:

$$p(x,y) = p(x)\cdot q(y).$$

In dat geval geldt:

$$q(y/x) = q(y)$$

$$p(x/y) = p(x).$$

Door middel van conditionele kansen kan men aangeven dat er afhankelijkheid bestaat tussen twee continue stochastische variabelen. In plaats daarvan wordt ook vaak de *covariantie* of *correlatie* gebruikt. Laat **x** en **y** twee continue stochastische variabelen zijn, dan wordt de *covariantie* gedefinieerd door:

$$\mathrm{cov}(\mathbf{x},\mathbf{y}) = E\big[(\mathbf{x} - E(x))\cdot(\mathbf{y} - E(\mathbf{y}))\big]$$

$$= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (x - E(x))\cdot(y - E(\mathbf{y}))\cdot p(x,y)\,\mathrm{d}x\mathrm{d}y. \tag{5.6}$$

Merk op dat de covariantie van de stochastische variabele **x** met zichzelf is:

$$\mathrm{cov}(\mathbf{x},\mathbf{x}) = E\big[(\mathbf{x} - E(x))^2\big] = \mathrm{var}(x).$$

Indien men de covariantie normeert ten opzichte van de varianties van **x** en **y** dan ontstaat de *correlatiecoëfficient* $\rho$, die we ook al tegenkwamen in de uitdrukking voor de twee-dimensionale gaussische verdeling. De correlatiecoëfficient is gegeven door:

$$\rho = \frac{\mathrm{cov}(\mathbf{x},\mathbf{y})}{\sqrt{\mathrm{var}(x)\cdot\mathrm{var}(\mathbf{y})}}. \tag{5.7}$$

Men kan bewijzen dat $|\rho| \leq 1$.
Een ander belangrijk begrip is het begrip *correlatie*. De *correlatie* tussen **x** en **y** is gedefinieerd door:

$$R(\mathbf{x},\mathbf{y}) = E(\mathbf{x}\cdot\mathbf{y}) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x\cdot y\cdot p(x,y)\,\mathrm{d}x\mathrm{d}y. \tag{5.8}$$

Het wordt aan de lezer over gelaten te bewijzen dat het volgende verband bestaat tussen covariantie en correlatie:

$$\mathrm{cov}(\mathbf{x},\mathbf{y}) = R(\mathbf{x},\mathbf{y}) - E(\mathbf{x})\cdot E(\mathbf{y}).$$

## 5.2. Stochastische signalen

Signalen waarvan hun waarde op een willekeurig tijdstip onvoorspelbaar is, worden stochastische signalen genoemd. Formeel definiëren we een stochastisch signaal als $\{\mathbf{x}(t),\ t \in T\}$, waarin $T$ een zeker tijdinterval kan zijn, maar ook een verzameling tijdstippen (bijvoorbeeld het begin van elk uur). Met $\mathbf{x}(t)$ wordt dan bedoeld de (toevallige) waarde van het signaal op het tijdstip $t$.

Op een bepaald tijdstip $t_0$ is $\mathbf{x}(t_0)$ een stochastische variabele die aangeeft welke waarden het signaal op tijdstip $t_0$ kan aannemen en hoe waarschijnlijk het is dat deze waarden inderdaad optreden. Omdat het signaal $\mathbf{x}(t)$ wordt gegeven voor alle $t$ uit een verzameling $T$ is hiermee in beginsel vastgelegd welke vormen het signaal $\mathbf{x}(t)$ als geheel kan aannemen en hoe waarschijnlijk het is dat zij voorkomen. Men kan dus een stochastisch signaal $\mathbf{x}(\mathrm{t})$ voor elk tijdstip $t \in T$ opvatten als een stochastische variabele. Anderzijds kan men ook voor alle tijdstippen uit de verzameling T nagaan hoe het signaal er telkens uitziet. Men beschouwt dan de gerealiseerde waarden van het signaal als één geheel. Men spreekt dan van een *realisatie* van het stochastische

signaal, aangegeven met $x(t)$. Een realisatie van een stochastisch signaal kan men als het ware opvaten als een uitkomst van het experiment: wek eenmalig het stochastisch signaal $\{\mathbf{x}(t),\ t \in T\}$ op en neem de uitkomst (dat wil zeggen de opeenvolgende signaalwaarden $[x(t),\ t \in T]$) waar.
Bij analoge signalen is de signaalwaarde op elk tijdstip een continue grootheid, zodat het stochastisch proces $\{\mathbf{x}(t),\ t \in T\}$ voor een bepaald tijdstip $t$ een continue stochastische variabele is met een bepaalde kansdichtheid $p(x)$. Ook is het signaal als functie van de tijd een continu verschijnsel, hetgeen inhoudt dat de verzameling $T$ een zeker tijdinterval is. Veelal is dit de meettijd of waarnemingstijd, maar het kan ook de intrinsieke tijd zijn dat het signaal optreedt.
Overigens moet worden bedacht dat in de voorgaande bespreking stilzwijgend is uitgegaan van een-dimensionale signalen als functie van de tijd. Men kan echter ook twee- en zelfs meer-dimensionale signalen beschouwen, bijvoorbeeld televisiebeelden of meerkanaals registraties van seismische signalen. Ook kan men de helderheid van een beeldscherm opvatten als een stochastisch proces waarbij de $(x,y)$-coördinaten van de beeldpunten op het scherm eenzelfde rol vervullen als de tijdparameter.
Hoewel een analoog signaal zowel wat betreft de signaalwaarden als de tijdparameter een continu karakter bezit, is dit een te complexe beschrijving om functioneel te zijn. We moeten ons veelal beperken tot zogenaamde tijddiscrete signalen die slechts op bepaalde tijdstippen bekend zijn. Alhoewel dit een beperking lijkt, blijkt dit doorgaans toch niet het geval te zijn, namelijk in die gevallen dat het stochastisch signaal een beperkte bandbreedte bezit.
Dit verband tussen de bandbreedte en de tijdsduur tussen de opeenvolgende tijdstippen waarop een signaal dient te worden vastgelegd wordt gelegd door het *bemonsteringstheorema*.

Als een signaal $x(t)$ een bandbreedte $W$ Hz bezit dan is $x(t)$ volledig bepaald door bemonsteringswaarden die op tijdintervallen $1/2W$ sec uiteen liggen.

Het bewijs van dit theorema wordt hier niet gegeven. Van belang is om op te merken dat voor bandbegrensde signalen volstaan kan worden met een tijddiscrete voorstelling, ook als deze dus van oorsprong tijdcontinu zijn. De bemonsteringen van een tijddiscreet stochastisch signaal worden aangeduid met $\mathbf{x}(k/2W)$. Veelal wordt gemakshalve $2W = 1$ gesteld, zodat we als aanduiding $\mathbf{x}(k)$, $k = 0,1,2,\ldots$ gebruiken. Als de tijdsduur van het signaal $T$ is zal dit erin resulteren dat er $2W{\cdot}T$ bemonsteringen van het signaal beschikbaar zijn.

Het beschouwen van het tijddiscrete signaal maakt de bestudering eenvoudiger, zonder dat de resultaten aan algemeenheid hebben ingeboet.
Terugkerend naar stochastische signalen $\{\mathbf{x}(t),\ t \in T\}$ kunnen we deze na toepassing van het bemonsteringstheorema voorstellen als een rij (tijd-discrete) bemonsteringen, waarbij per bemonstering de waarde ervan als stochastische variabele kan worden beschouwd. Om het gehele stochastische signaal vast te leggen moet men dus voor de

totale verzameling van N bemonsteringstijdstippen de gezamenlijke kansdichtheid $p(x) = p(x_1,x_2,\ldots,x_N)$ vastleggen. Elke "waarde" van **x** is dan in feite een realisatie van het stochastisch proces $\{\mathbf{x}(t),\ t \in T\}$ vastgelegd op de bemonsteringstijdstippen $t_1,t_2,\ldots,t_N$.

In veel gevallen is het niet nodig of zelfs onmogelijk om de gezamenlijke kansdichtheid van alle bemonsteringen te gebruiken. Beperken we ons tot een eerste-orde beschrijving, dat wil zeggen met $N = 1$, dan kunnen we de signaaleigenschappen per bemonstering beschouwen. Omdat afhankelijkheid tussen bemonsteringen in de praktijk een grote rol speelt is doorgaans een tweede-orde beschrijving, dat wil zeggen met $N = 2$, gewenst of noodzakelijk. Doorgaans is dit ook voldoende. Een tweede-orde stochastisch signaal wordt dan vastgelegd met $\{\mathbf{x}(t_i,t_j),\ t_i,t_j \in T\}$ ofwel door middel van twee bemonsteringen of stochastische variabelen met gezamenlijke kansdichtheid $p(x_i,x_j)$.
In beginsel kan men voor een gegeven kansdichtheid $p(x_i,x_j)$ het signaal analyseren. Vaak beschouwt men echter niet de kansdichtheid zelf maar grootheden als de verwachting of gemiddelde waarde, de correlatie of de covariantie.
De autocorrelatiefunctie wordt gedefinieerd als

$$R_x(t_i,t_j) = E\{\mathbf{x}(t_i)\cdot\mathbf{x}(t_j)\} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} x_i x_j \cdot p(x_i,x_j)\,\mathrm{d}x_i\mathrm{d}x_j$$

en hangt dus af van de tijdstippen $t_i$ en $t_j$.
Nemen we nu voor de tijdstippen $t_i$ en $t_j$ de bemonsteringstijdstippen dan kunnen we de waarden van de autocorrelatiefunctie ook aangeven met behulp van een matrix, de zogenaamde autocorrelatiematrix $R_x$, waarvoor geldt

$$R_x = \begin{bmatrix} R_{11} & \ldots & R_{1N} \\ \vdots & & \vdots \\ R_{N1} & \ldots & R_{NN} \end{bmatrix},$$

waarin $R_{ij} = R(t_i,t_j)$ dus betrekking heeft op de correlatie van de bemonsteringen $\mathbf{x}_i$ en $\mathbf{x}_j$ op de tijdstippen $t_i$ en $t_j$ voor $i,j = 1,2,\ldots,N$.
Deze definitie is nog zeer algemeen omdat de correlatie tussen twee bemonsteringen $x_i$ en $x_j$ voor verschillende paren tijdstippen i en j verschillend mag zijn. Dergelijke signalen worden niet-stationair genoemd. Vaak kan men echter aannemen dat de signalen stationair zijn in die zin dat de correlatie uitsluitend afhangt van het tijdverschil $\tau = t_i - t_j$ maar niet van de absolute tijdstippen $t_i$ en $t_j$. Men spreekt dan van een zwak-stationair signaal, waarvoor dus geldt dat

$$R_{\mathrm{x}}(t_i,t_j) = R_{\mathrm{x}}(t_i - t_j) = R_{\mathrm{x}}(\tau) = E\{\mathbf{x}(t_i)\,\mathbf{x}(t_i - \tau)\}.$$

Omdat de correlatie slechts een afgeleide grootheid is van de gezamenlijke kansdichtheid $p(x_i,x_j)$ kan het voorkomen dat bij een zwak-stationair signaal de kansdichtheid $p(x_i,x_j)$ nog altijd wel afhangt van de absolute tijdstippen $t_i$ en $t_j$. Men gebruikt daarom ook het begrip strikt-stationair om aan te geven dat de kansdichtheid $p(x_i,x_j)$ zelf tijdinvariant is, dat wil zeggen alleen afhangt van $t_i - t_j$. Een strikt-stationair signaal is tevens zwak-stationair, maar het omgekeerde geldt dus niet altijd.
Naast het begrip correlatie wordt ook vaak gebruik gemaakt van de covariantie. Deze is eerder gedefinieerd als

$$K(x_i,x_j) = \int\int (x_i - E\{\mathbf{x}_\mathrm{i}\})\cdot(x_j - E\{\mathbf{x}_j\})\cdot p(x_i,x_j)\,\mathrm{d}x_i\mathrm{d}x_j,$$

waarin $E\{\mathbf{x}_i\}$ en $E\{\mathbf{x}_j\}$ de verwachtingen van $\mathbf{x}_i$ en $\mathbf{x}_j$ aanduiden. Het verschil met de correlatie is dus uitsluitend dat nu de verwachting van $x_i$ en $x_j$ wordt afgetrokken van $x_i$ en $x_j$. Bij veel toepassingen zijn deze verwachtingswaarden nul, zodat dan de correlatie en de covariantie aan elkaar gelijk zijn. De autocovariantiematrix voor $N$ bemonsteringen is

$$K = \begin{bmatrix} k_{11} & \dots & k_{1N} \\ \vdots & & \vdots \\ k_{N1} & \dots & k_{NN} \end{bmatrix}, \tag{5.9}$$

waarin $k_{ij} = K(x_i,x_j)$. Eenvoudig is in te zien dat de autocorrelatiematrix en de autocovariantiematrix symmetrisch zijn, dat wil zeggen dat geldt dat $R_{ij} = R_{ji}$ en $k_{ij} = k_{ji}$.

Een belangrijke klasse signalen, waarbij het voldoende is dat het signaal zwak stationair is, wordt gevormd door gaussische stochastische signalen. Voor deze signalen is de autocovariantiematrix een voldoende beschrijving van het signaal.

Een stochastisch signaal $\{\mathbf{x}(t), t \in T\}$ heet een *Gaussisch signaal* als al zijn $N$-dimensionale kansdichtheden $p(x_1,\dots,x_N)$ voor $N = 1,2,\dots$ $N$-dimensionale gaussische kansdichtheden zijn.
Voor $N = 1$, dus voor één bemonstering, is de gaussische kansdichtheid (zie ook formule (5.3)):

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{(x-\mu)^2}{2\sigma^2}\right\}$$

waarin $\mu$ de gemiddelde waarde is en $\sigma^2$ de variantie. Voor $N = 2$ is de tweedimensionale gaussische verdeling gegeven in formule (5.4).

In het algemene geval van een $N$-dimensionale gaussische verdeling is het eenvoudig om een vectornotatie te hanteren. De uitdrukking voor $p(\tilde{\mathbf{x}})$ is dan

$$p(\tilde{\mathbf{x}}) = \frac{1}{(2\pi)^{N/2}|K_\mathrm{x}|^{1/2}} \exp\{-\tfrac{1}{2}(\tilde{x}-\tilde{\mu})\cdot K_\mathrm{x}^{-1}(\tilde{x}-\tilde{\mu})^\mathrm{T}\}, \tag{5.10}$$

waarin $\tilde{\mu}$ de vector van de gemiddelde waarden is, $K_x$ de autocovariantiematrix van het proces en $|K_x|$ de determinant van $K_x$.
Het gaussische signaal is het enige signaal dat volledig beschreven wordt door zijn gemiddelde en zijn covariantiematrix.

Als het proces stationair is, is de covariantie $k_{ij}$ alleen afhankelijk van het tijdverschil $|i–j|$ tussen $x_i$ en $x_j$. Dit betekent bijvoorbeeld dat voor alle tijdstippen $t_i$ en $t_j$ zodanig dat $j = i + 1$ of $j = i - 1$ geldt dat $k_{ij} = k_1$. De autocovariantiematrix heeft dan de volgende gedaante

$$K_x = \begin{bmatrix} k_0 & k_1 & \cdot & \cdot & \cdot & k_{N-1} \\ k_1 & k_0 & \cdot & & & \cdot \\ \cdot & & \cdot & \cdot & & \cdot \\ \cdot & & & \cdot & \cdot & k_1 \\ k_{N-1} & & \cdot & \cdot & k_1 & k_0 \end{bmatrix} \tag{5.11}$$

en vertoont dus een zeer specifieke structuur. Een matrix van deze vorm wordt een Toeplitz-matrix genoemd. De eerste rij van deze matrix geeft juist de bemonsterde waarden aan van de autocovariantiefunctie van het gaussisch proces.
Een gaussisch signaal heeft een aantal eigenschappen die het gebruik ervan aantrekkelijk maken. Wellicht de belangrijkste eigenschap is dat een lineaire bewerking die op een gaussisch signaal wordt uitgevoerd leidt tot een signaal dat eveneens gaussisch is. Hierdoor speelt het gaussisch proces bij de analyse van stochastische signalen een rol die vergelijkbaar is met de rol van lineaire systemen in de systeemtheorie.
De beschrijving van stochastische signalen is tot dusver beperkt gebleven tot het tijddomein. Men kan echter ook het frequentiedomein hiervoor gebruiken. Van groot belang is dan het vermogensdichtheidsspectrum.
Dit vermogensdichtheidsspectrum $S(\omega)$ veelal kortheidshalve het spectrum genoemd, kan men vinden uit de autocorrelatiefunctie met behulp van een Fouriertransformatie. Het is als volgt gedefinieerd:

$$S(\omega) = \int_{-\infty}^{\infty} R_x(\tau)\cdot e^{-j\omega\tau}\,\mathrm{d}\tau, \tag{5.12}$$

$$R(\tau) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega)\cdot e^{+j\omega\tau}\,\mathrm{d}\omega. \tag{5.13}$$

In het geval $\tau = 0$ vindt men

$$R(0) = \frac{1}{2\pi}\int_{-\infty}^{\infty} S_x(\omega)\,\mathrm{d}\omega = P_x. \tag{5.14}$$

Dit betekent dat het gemiddelde vermogen van een signaal kan worden gevonden uit de autocorrelatiefunctie door $\tau = 0$ te stellen, maar ook door het vermogensdichtheids-spectrum te integreren over het gehele frequentiebereik.

## 5.3. Het continue informatiebegrip

Bij het behandelen van het discrete informatiebegrip is naar voren gekomen dat de informatiemaat nauw samenhangt met de kansen van optreden van de verschillende gebeurtenissen ofwel met de kansverdeling van een discrete stochastische variabele. Ook voor een continu stochastisch signaal, waarbij de bemonsteringen een continu waardenbereik bezitten, kan een informatiemaat worden afgeleid. In de vorige paragraaf werd reeds aandacht besteed aan aan continue stochastische variabelen gerelateerde kansdichtheden. Naar analogie met het discrete geval komen we tot de volgende definitie van de marginale informatiemaat.

### *Definitie 5.1*

Voor een continue stochastische variabele **x** met kansdichtheid $p(x)$ is de hoeveelheid informatie gelijk aan:

$$H(X) = -\int_{-\infty}^{\infty} p(x) \cdot \log p(x)\, \mathrm{d}x. \tag{5.15}$$

□

De op deze wijze voor een continue stochastische variabele gedefinieerde $H(X)$ kan negatief worden, zulks in tegenstelling tot het discrete geval. Dit is niet in overeenstemming met het intuïtieve informatiebegrip. Deze situatie doet zich voor indien de kansdichtheid $p(x)$ groter is dan 1. De oorzaak hiervan is dat geen rekening is gehouden met het feit dat men in praktische toepassingen altijd te maken heeft met een eindige meetnauwkeurigheid (d.w.z. de variabele **x** heeft een minder fijne onderverdeling), zodat het geval $H(X) < 0$ als onrealistisch beschouwd kan worden.
Ook voor een continue stochastische variabele is het van belang om na te gaan welke kansdichtheid leidt tot een maximale hoeveelheid informatie. De afleiding is echter verschillend van de afleiding die in het discrete geval is gegeven. Dit komt voort uit het feit dat bij een continue stochastische variabele doorgaans extra restricties moeten worden opgelegd om te voorkomen dat $H(X)$ oneindig groot wordt. Deze restricties kunnen bijvoorbeeld zijn een begrensde amplitude of een constant vermogen (variantie). De aard van deze restricties is dan mede bepalend voor de aard van de kansdichtheid die leidt tot een maximale hoeveelheid informatie. Er zullen twee gevallen beschouwd worden, namelijk amplitude begrenzing, en begrenzing van het vermogen (ofwel de variantie). In de volgende stelling wordt nagegaan welke kansdichtheid leidt tot een maximale hoeveelheid informatie als het waardenbereik begrensd wordt tot $-A$ en $+A$.

### Stelling 5.1

Voor een amplitude begrensd signaal met waardenbereik $(-A, +A)$ geldt dat $H(X)$

maximaal is dan en slechts dan als

$$p(x) = \frac{1}{2A}\,.$$

De maximale waarde is gegeven door:

$$H(X) = \log 2A.$$

### *Bewijs*

Om dit probleem op te lossen wordt gebruik gemaakt van een methode uit de variatie-rekening. De opgave is om de kansdichtheid $p(x)$ te bepalen waarvoor:

$$H(X) = -\int_{-\infty}^{\infty} p(x)\cdot\log p(x)\,\mathrm{d}x$$

maximaal is, waarbij $p(x)$ vanwege de amplitudebegrenzing moet voldoen aan:

$$\int_{-A}^{A} p(x)\,\mathrm{d}x = 1.$$

Hiertoe wordt de functie:

$$G(x) = -p(x)\cdot\log p(x) + \alpha\cdot p(x)$$

gedifferentieerd naar $p(x)$ en vervolgens gelijk aan nul gesteld. Dit geeft:

$$-\log p(x) - \log e + \alpha = 0$$

ofwel:

$$\ln p(x) = \frac{\alpha}{\log e} - 1 = k, \quad \text{zodat} \quad p(x) = e^k.$$

Nu geldt: $\int_{-A}^{A} p(x)\,\mathrm{d}x = 1$, zodat substitutie van $p(x)$ geeft:

$$\int_{-A}^{A} e^k\,\mathrm{d}x = 1 \quad \Rightarrow \quad \left[e^k\cdot x\right]_{-A}^{A} = e^k\cdot 2A = 1,$$

zodat:

$$p(x) = \frac{1}{2A}\,.$$

Substitutie van $p(x) = \dfrac{1}{2A}$ in $H(X)$ levert tenslotte:

$$H(X) = -\int_{-A}^{A} \frac{1}{2A}\log\left(\frac{1}{2A}\right)\mathrm{d}x = \log 2A. \qquad \square$$

Het blijkt dus dat de uniforme kansdichtheid in het geval van amplitudebegrenzing een maximale hoeveelheid informatie levert evenredig aan de maximale amplitude. Dit is enigszins in overeenstemming met het geval van een discrete stochastische variabele, die immers een maximale hoeveelheid informatie bezit als deze een symmetrische kansverdeling heeft.

Een ander belangrijk geval is het begrenzen van het vermogen van een signaal, hetgeen neerkomt op het vastleggen van de variantie van de bemonsteringen.

**Stelling 5.2**

Voor een signaal met constant vermogen $\sigma^2$,

$$\sigma^2 = \int_{-\infty}^{\infty} x^2 p(x)\,\mathrm{d}x,$$

geldt dat $H(X)$ maximaal is dan en slechts dan als:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}.$$

De maximale hoeveelheid informatie die hierbij hoort is

$$H(X) = \log(\sigma\sqrt{2\pi e}).$$

***Bewijs***

We moeten nu de kansdichtheid $p(x)$ bepalen, zodanig dat:

$$H(X) = -\int_{-\infty}^{\infty} p(x)\cdot\log p(x)\,\mathrm{dx}$$

maximaal is, waarbij voldaan moet zijn aan de restricties:

$$\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = 1$$

en:

$$\int_{-\infty}^{\infty} x^2 p(x)\,\mathrm{d}x = \sigma^2$$

waarbij dus $\sigma^2$ constant wordt verondersteld. We vormen nu de functie:

$$G(x) = -p(x)\log p(x) + \alpha_1 p(x) + \alpha_2 x^2 p(x)$$

en stellen de afgeleide van $G(x)$ naar $p(x)$ gelijk aan nul. Dit geeft:

$$-\log p(x) - \log e + \alpha_1 + \alpha_2 x^2 = 0$$

ofwel na delen door $-{}^2\log e$:

$$\ln p(x) + 1 - \lambda_1 - \lambda_2 \cdot x^2 = 0$$

met: $$\lambda_1 = \frac{\alpha_1}{\log e} \quad \text{en} \quad \lambda_2 = \frac{\alpha_2}{\log e}.$$

Dit geeft als oplossing:

$$p(x) = e^{\lambda_1 - 1} \cdot e^{\lambda_2 x^2}.$$

De parameters $\lambda_1$ en $\lambda_2$ worden geëlimineerd door substitutie van $p(x)$ in de beide restricties:

$$\int_{-\infty}^{\infty} p(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} e^{\lambda_1 - 1} \cdot e^{\lambda_2 x^2}\,\mathrm{d}x = 1.$$

Dit geeft:

$$e^{\lambda_1 - 1} = \sqrt{-\frac{\lambda_2}{\pi}}.$$

Verder geldt:

$$\int_{-\infty}^{\infty} x^2 p(x)\,\mathrm{d}x = \int_{-\infty}^{\infty} x^2 e^{\lambda_2 x^2} \sqrt{-\frac{\lambda_2}{\pi}}\,\mathrm{d}x = \sigma^2,$$

waaruit volgt:

$$\lambda_2 = -\frac{1}{2\sigma^2}$$

en dus: $$e^{\lambda_1 - 1} = \frac{1}{\sigma\sqrt{2\pi}}.$$

Dit geeft tenslotte:

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\}.$$

De hoeveelheid informatie die hierbij hoort is:

$$H(X) = -\int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} \cdot \log \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} \mathrm{d}x =$$

$$= \log \sigma\sqrt{2\pi} \cdot \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} \mathrm{d}x +$$

$$+\int_{-\infty}^{\infty} \frac{x^2 \log e}{2\sigma^2} \cdot \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{x^2}{2\sigma^2}\right\} \mathrm{d}x =$$

$$= \log \sigma\sqrt{2\pi} + \frac{\log e}{2\sigma^2} \cdot \mathrm{var}(x) = \log \sigma\sqrt{2\pi} + \tfrac{1}{2}\log e =$$

$$= \log \sigma\sqrt{2\pi e}\,. \qquad \square$$

Het blijkt dus dat bij een constant vermogen $\sigma^2$ de normale verdeling een maximale hoeveelheid informatie levert. $H(X)$ is evenredig met de logaritme van de standaard deviatie $\sigma$. Juist omdat in technische toepassingen zowel het vermogen als de normale verdeling veelvuldig worden gebruikt is dit een belangrijk resultaat.
Met betrekking tot het discrete geval hebben we naast de marginale informatiemaat ook de conditionele, gezamenlijke en wederzijdse informatiematen gedefinieerd. Voor het continue geval gaat dit evenzo. Hieronder zijn in het kort de conditionele, gezamenlijke en wederzijdse informatiematen voor het continue geval geïntroduceerd. Merk op dat de onderlinge relaties tussen de diverse maten in grote lijnen overeenstemmen met die voor het discrete geval.
Is er sprake van een aantal bemonsteringen van een signaal, waarbij elke bemonstering een continue stochastische variabele is, dan kan men een gezamenlijke hoeveelheid informatie definiëren. Hierbij wordt uitgegaan van de gezamenlijke kansdichtheid. Voor het geval dat er twee stochastische variabelen **x** en *y* zijn met gezamelijke kansdichtheid $p(x,y)$ is de *gezamenlijke hoeveelheid informatie* gedefinieerd als:

$$H(X,Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\cdot\log p(x,y)\,\mathrm{d}x\mathrm{d}y.$$

De conditionele hoeveelheid informatie kan als volgt gevonden worden. Voor twee continue stochastische variabelen **x** en **y** kan men de gezamenlijke kansdichtheid schrijven als:

$$p(x,y) = p(x)\cdot q(y/x) = q(y)\cdot p(x/y).$$

We definiëren nu de *conditionele hoeveelheid informatie* van $x$ gegeven $y$ als:

$$H(X/Y) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\cdot\log p(x/y)\,\mathrm{d}x\mathrm{d}y.$$

Evenzo: $$H(Y/X) = -\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} p(x,y)\cdot\log q(y/x)\,\mathrm{d}x\mathrm{d}y.$$

Deze definities komen dus sterk overeen met de reeds eerder voor het discrete geval gegeven definities. Evenals voor het discrete geval kan men afleiden dat geldt:

$$H(X/Y) \leq H(X)$$

$$H(Y/X) \leq H(Y),$$

met gelijkheid als **x** en **y** stochastische onafhankelijk zijn. Tevens geldt:

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y).$$

De bewijzen zullen hier niet worden gegeven maar verlopen identiek aan de bewijzen voor het discrete geval. Ook geldt op grond van het voorgaande, dat:

$$H(X,Y) \leq H(X) + H(Y),$$

met gelijkheid als **x** en **y** stochastisch onafhankelijk zijn. Men kan echter *niet* stellen dat $H(Y/X) \geq 0$ is. Definiëren we tenslotte de wederzijdse hoeveelheid informatie als

$$I(X;Y) = H(X) + H(Y) - H(X,Y),$$

dan vinden we met behulp van eerder gegeven definities dat $I(X;Y)$ gelijk is aan:

$$I(X;Y) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} p(x,y) \cdot \log \frac{p(x,y)}{p(x) \cdot q(y)} \,\mathrm{d}x\mathrm{d}y.$$

## 5.4. Opgaven

**5.1.** *a.* De stochastische variabele **x** representeert de amplitude van een signaal $\mathbf{x}(t)$, dat begrensd is tussen –3 en +3 volt. Tussen deze grenzen heeft **x** een uniforme verdeling. Bepaal de hoeveelheid informatie $H(X)$.

*b.* Bepaal de hoeveelheid informatie $H(X)$ als **x** uniform verdeeld is tussen –5 en +5 volt.

*c.* Licht het verschil tussen de in a. en b. gevonden antwoorden toe.

**5.2.** Een bemonstering **x** van een stochastisch signaal is uniform verdeeld tussen +1 en +7 volt.

*a.* Bepaal de hoeveelheid informatie $H(X)$. Welke conclusies kunt u trekken als u dit resultaat vergelijkt met dat van opgave 5.1.a?

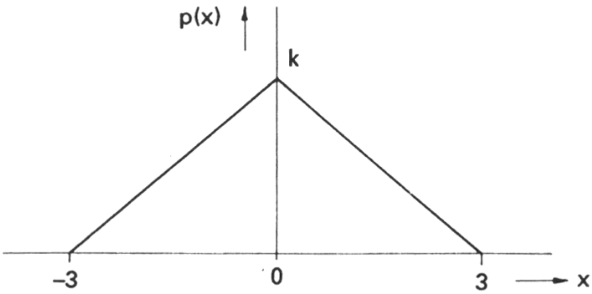
*b.* Bepaal $E(x)$ en var $(x)$.

**5.3.** Gegeven is de continue stochastische variabele **x** met kansdichtheid $p_{\mathrm{X}}(x)$. Voor de stochastische variabele $y$ geldt:

$$y = \mathbf{x} + \alpha.$$

Bewijs dat geldt: $H(X) = H(Y)$.

**5.4.** De stochastische variabele **x** heeft een kansdichtheid als gegeven in onderstaande figuur.




*a.* Bepaal de waarde van *k*.

*b.* Bepaal *H*(*X*).

*c.* Vergelijk de gevonden waarde van *H*(*X*) met het antwoord op opgave 5.1.a en licht een eventueel verschil toe.

**5.5.** Een bemonstering **x** van het stochastische signaal **x**(*t*) is uniform verdeeld tussen –*A* en +*A* met *A* > 0.

*a.* Schets *H*(*X*) als functie van *A*.

*b.* Bepaal en schets var(*x*) als functie van *A*.

*c.* Wat is de overeenkomst tussen beide grafieken?

**5.6.** De stochastische variabele **x** heeft een negatief-exponentiële verdeling:

$$p(x) = \frac{1}{\lambda}\exp(-\frac{x}{\lambda}) \qquad \mathrm{x} \geq 0,$$

$$= 0, \qquad x > 0.$$

*a.* Bereken de hoeveelheid informatie *H*(*X*).

*b.* Beschouw een stochastische variabele die geen negatieve waarden kan aannemen en waarvan de verwachting $\lambda$ is. Toon aan dat deze variabele een maximale hoeveelheid informatie bezit indien zijn kansdichtheid de hierboven gegeven verdeling is.

**5.7.** Een bemonstering **x** van een signaal *x*(*t*) heeft de volgende kansdichtheid:

$$p(x) = \frac{1}{2\lambda}\exp(-\frac{|x|}{\lambda}).$$

Het signaal *x*(*t*) wordt dubbelfasig gelijkgericht, zodat voor een bemonstering **y** van het uitgangssignaal geldt:

$$\mathbf{y} = |\mathbf{x}|.$$

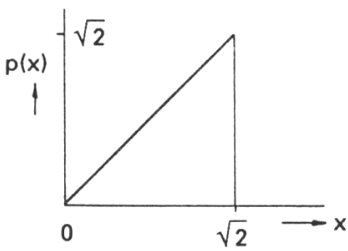
*a.* Bepaal *H*(*X*).

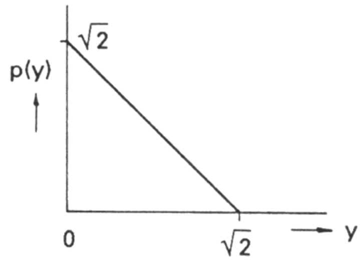
*b.* Bepaal *q*(*y*) en met behulp daarvan *H*(*Y*).

*c.* Beredeneer dat $H(Y/X) = 0$.

*d.* Bepaal uit de vorige antwoorden $H(X/Y)$. Kunt u de gevonden waarde verklaren?

**5.8.** De stochastische onafhankelijke variabelen $\mathbf{x}$ en $\mathbf{y}$, die signalen $x(t)$ en $y(t)$ representeren, hebben de volgende kansdichtheden:

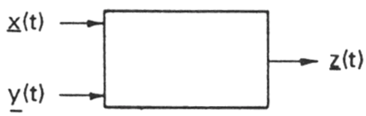



De signalen $\mathbf{x}(t)$ en $\mathbf{y}(t)$ worden toegevoerd aan een discriminator, waarbij voor een bemonstering $\underline{z}$ van het uitgangssignaal geldt:

$$\mathbf{z} = \mathbf{x} \qquad \text{voor} \qquad \mathbf{x} \geq \frac{1}{2}\sqrt{2}$$

$$\mathbf{z} = \underline{y} \qquad \text{voor} \qquad \mathbf{x} < \frac{1}{2}\sqrt{2}\,.$$



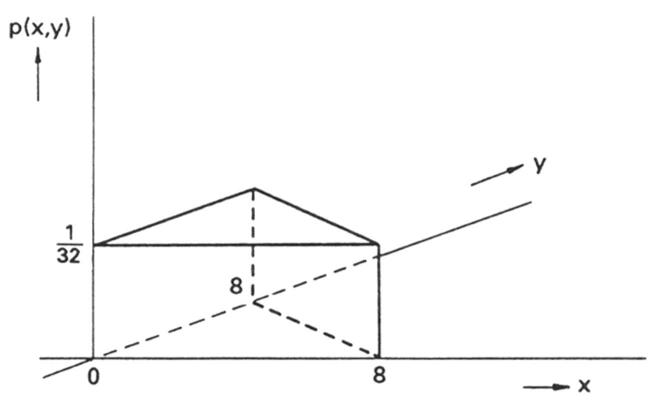

*a.* Bepaal $H(Z/x < \frac{1}{2}\sqrt{2})$.

*b.* Bepaal $H(Z/x \geq \frac{1}{2}\sqrt{2})$.

*c.* Bepaal $H(Z)$.

Aanwijzing: bij b. dient u te bedenken dat $p(x) \neq p(x/x \geq \frac{1}{2}\sqrt{2})$.

**5.9.** De stochastische variabelen $\mathbf{x}$ en $\mathbf{y}$ hebben de hier geschetste gezamenlijke kansdichtheid:



*a.* Bepaal $H(X)$.
*b.* Bepaal $H(Y/X)$.
*c.* Bepaal uit a. en b. de grootheid $H(X/Y)$.
Controleer uw antwoord door directe berekening.
*d.* Concludeer op grond van uw antwoord of **x** en **y** stochastisch onafhankelijk zijn.

**5.10.** Van de stochastische variabelen **x** en **y** is gegeven:

- $p(x,y)$ is de twee-dimensionale gaussiche kansdichtheid;
- $E(\mathbf{x}) = E(\mathbf{y}) = 0$;
- var $(\mathbf{x}) = \sigma_1^2$ en var $\mathbf{y}) = \sigma_2^2$;
- de correlatiecoëfficiënt is $\rho$.

*a.* Bepaal $H(Y)$.
*b.* Bepaal $H(X/Y)$.
*c.* Bepaal $H(X,Y)$.
*d.* Hoe worden deze grootheden beïnvloed als $\rho = 0$ en hoe verklaart u dit?

## 5.5. Uitwerkingen

**5.1.** *a.* In het algemeen geldt voor een stochastische variabele **x** die tussen $-A$ en $+A$ begrensd is, terwijl **x** tussen deze grenzen een uniforme verdeling heeft, dat:

$$p(x) = \frac{1}{2A} \quad \text{voor} \quad -A \leq x \leq A.$$

De corresponderende hoeveelheid informatie is op basis van stelling 5.1 gelijk aan:

$$H(X) \;\; = \log 2A.$$

In het onderhavige geval geldt $A = 3$, dus:

$$H(X) = \log 6 = 2{,}58 \;\; \text{bit.}$$

*b.* Substitutie van $A = 5$ in de onder a. gegeven algemene formule leidt tot

$$H(X) = \log 10 = 3{,}32 \;\; \text{bit.}$$

*c.* Er kan geconcludeerd worden dat de hoeveelheid informatie zal toenemen als het bereik van de variabele groter is. Dit stemt overeen met het feit dat de onzekerheid omtrent de variabele toeneemt bij een groter bereik.

**5.2.** *a.* Aangezien er sprake is van een uniforme verdeling geldt $p(x) = \frac{1}{6}$ voor $1 \leq x \leq 7$. De hoeveelheid informatie is gelijk aan

$$H(X) = -\int_1^7 p(x)\cdot\log p(x)\ \mathrm{d}x = -\int_1^7 \tfrac{1}{6}\ \log\tfrac{1}{2}\ \mathrm{d}x = \log 6 = 2{,}58\ \text{ bit.}$$

Vergelijking met het resultaat van opgave 5.1.a leert dat de hoeveelheid informatie bij een uniforme verdeling alleen afhangt van het bereik van de variabele, en niet van de ligging van het bereik op de getallenrechte.

*b.* Er geldt:

$$E(x) = \int_1^7 x\cdot p(x)\ \mathrm{d}x = \int_1^7 \tfrac{x}{6}\ \mathrm{d}x = \tfrac{1}{2}\,x^2\Big|_1^7 = 4.$$

Voor de variantie geldt: $\mathrm{var}(x) = E[(x - E(x))^2]$.

Aangezien $E(x) = 4$ volgt:

$$\mathrm{var}(x) = E[(x-4)^2] = \int_1^7 (x-4)^2 p(x)\ \mathrm{dx} =$$

$$= \int_1^7 \frac{(x-4)^2}{6}\ \mathrm{d}x = \int_{-3}^{3} \frac{y^2}{6}\ \mathrm{d}y = \frac{1}{18}\,y^3\Big|_{-3}^{3} = 3.$$

**5.3.** Aangezien $\alpha$ een constante is, zal het toevoegen van $\alpha$ aan $\mathbf{x}$ de onzekerheid in $\mathbf{y} = \mathbf{x} + \alpha$ niet wijzigen. De hoeveelheid informatie zal dan ook ongewijzigd blijven en er zal gelden: $H(X) = H(Y)$. Het is ook in te zien door te bedenken dat de onzekerheid met betrekking tot $\mathbf{y}$ feitelijk gelijk is aan de onzekerheid met betrekking tot $\mathbf{x}$ en $\alpha$: $H(X,\alpha)$.

Aangezien $\alpha$ een constante is waarvan de waarde a priori vastligt, zal gelden $H(\alpha) = 0$. Er volgt nu:

$$H(Y) = H(X,\alpha) = H(X) + H(\alpha) = H(X).$$

**5.4.** *a.* Er moet gelden

$$\int_{-\infty}^{\infty} p(x)\ \mathrm{d}x = 1.$$

Eenvoudig is in te zien, dan dan $k = \frac{1}{3}$ moet zijn.

*b.* Teneinde $H(X)$ te kunnen uitrekenen, dient eerst $p(x)$ gevonden te worden. Er kan nagegaan worden dat

$$p(x) = \frac{x+3}{9} \qquad \text{voor } x \leq 0,$$

en

$$p(x) = \frac{3-x}{9} \qquad \text{voor } x > 0.$$

Voor $H(X)$ leidt dit tot

$$H(X) = -\int_{-3}^{0} \frac{x+3}{9} \log\left(\frac{x+3}{9}\right) dx - \int_{0}^{3} \frac{3-x}{9} \log\left(\frac{3-x}{9}\right) dx =$$

$$= -18 \int_{0}^{1/3} y \cdot \log y \, dy.$$

Aangezien algemeen geldt

$$\int y \cdot \ln y \, dy = y^2\left[\tfrac{1}{2} \ln y - \tfrac{1}{4}\right],$$

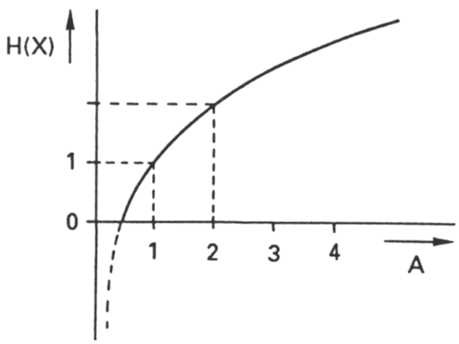
volgt voor $H(X)$:

$$H(X) = -\frac{18}{\ln 2} \int_{0}^{1/3} y \cdot \ln y \, dy = -\frac{18}{\ln 2} y^2\left(\tfrac{1}{2} \ln y - \tfrac{1}{4}\right)\Big|_{0}^{1/3} =$$

$$= -\log \tfrac{1}{3} + \tfrac{1}{2} \log e \approx 2{,}30 \text{ bit.}$$

*c.* De hier gevonden waarde is lager dan die in opgave 5.1.a. Dit hangt samen met het feit dat in de onderhavige opgave de kansdichtheid een piek vertoont, waardoor de onzekerheid en daarmee de hoeveelheid informatie afneemt.

**5.5.** *a.* Met opgave 5.1.a volgt $H(X) = \log 2A$. De grafiek is gegeven in onderstaande figuur.

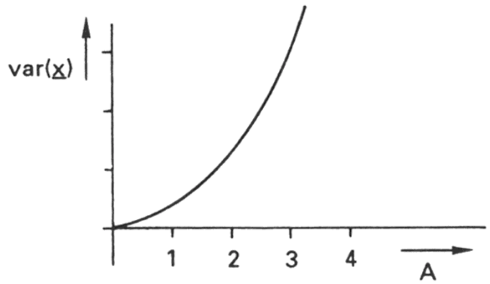



*b.* Var($x$) kan als volgt berekend worden:

$$E(x) = \int_{-A}^{A} x \cdot p(x) \, dx = \int_{-A}^{A} \frac{1}{2A} x \, dx = \frac{1}{4A} x^2 \Big|_{-A}^{A} = 0.$$

Hiermee wordt var($x$) gelijk aan

$$\text{var}(x) = E[(x - E(x))^2] = E[x^2] = \int_{-A}^{A} x^2 \cdot p(x) \, dx = \frac{1}{6A} x^3 \Big|_{-A}^{A} = \tfrac{1}{3} A^2.$$




*c*. Hoe groter $A$ des te groter de variantie en des te groter de onzekerheid omtrent **x**. Dit laatste correspondeert met de toename van $H(X)$ zoals gegeven in de grafiek van $H(X)$ als functie van $A$.

**5.6.** a.
$$H(X) = -\int_0^\infty p(x)\cdot\log p(x)\,\mathrm{d}x = -\log e \int_0^\infty p(x)\cdot\ln p(x)\,\mathrm{d}x =$$
$$= -\log e \int_0^\infty \frac{1}{\lambda}\exp(-\frac{x}{\lambda})\cdot\ln[\frac{1}{\lambda}\exp(-\frac{x}{\lambda})]\mathrm{d}x =$$
$$= \log e\cdot\frac{\ln\lambda}{\lambda}\int_0^\infty \exp(-\frac{x}{\lambda})\,\mathrm{d}x + \log e\,\frac{1}{\lambda^2}\int_0^\infty x\exp(-\frac{x}{\lambda})\,\mathrm{d}x =$$
$$= \log e\cdot\ln\lambda\int_0^\infty e^{-y}\mathrm{d}y + \log e\int_0^\infty ye^{-y}\mathrm{d}y =$$
$$= -\log\lambda e^{-y}\Big|_0^\infty + \log e\, e^{-y}(-y-1)\Big|_0^\infty = \log\lambda + \log e.$$

b. Er dient een kansdichtheid $p(x)$ bepaald te worden zodanig dat

$$H(X) = -\int_0^\infty p(x)\cdot\log p(x)\,\mathrm{d}x$$

maximaal is, waarbij voldaan moet zijn aan

$$\int_0^\infty p(x)\,\mathrm{d}x = 1 \qquad \text{en} \qquad \int_0^\infty x\cdot p(x)\,\mathrm{d}x = \lambda.$$

We vormen nu de functie:

$$G(x) = -p(x)\cdot\log p(x) + \alpha_1\cdot p(x) + \alpha_2\cdot x\cdot p(x).$$

Stellen we de afgeleide van $G(x)$ naar $p(x)$ gelijk aan nul dan vinden we:

$$-{}^2\log p(x) - {}^2\log e + \alpha_1 + \alpha_2\cdot x = 0.$$

Deze laatste uitdrukking gaat na delen door $-{}^2\log e$ over in

$$\ln p(x) + 1 - \lambda_1 - \lambda_2\cdot x = 0,$$

waarbij $\lambda_1 = \alpha_1/{}^2\log e$ en $\lambda_2 = \alpha_2/{}^2\log e$. Dit geeft als oplossing:

$$p(x) = e^{\lambda_1-1} \cdot e^{\lambda_2 x}.$$

Substitutie van deze $p(x)$ in de beide randvoorwaarden levert:

$$\int_0^\infty p(x)\,dx = \int_0^\infty e^{\lambda_1-1} \cdot e^{\lambda_2 x}\,dx = \left[e^{\lambda_1-1} \cdot e^{\lambda_2 x}/\lambda_2\right]_0^\infty = -e^{\lambda_1-1}/\lambda_2 = 1,$$

waarbij aangenomen is dat $\lambda_2 < 0$. Dus geldt: $e^{\lambda_1-1} = -\lambda_2$.
Ook moet gelden

$$\int_0^\infty x \cdot p(x)\,dx = \int_0^\infty x \cdot e^{\lambda_1-1} \cdot e^{\lambda_2 x}\,dx = -\int_0^\infty x \cdot \lambda_2 \cdot e^{\lambda_2 x}\,dx = \lambda,$$

waaruit volgt

$$\lambda = -\int_0^\infty x \cdot \lambda_2 \cdot e^{\lambda_2 x}\,dx = -\lambda_2 \frac{e^{\lambda_2 x}}{(\lambda_2)^2}(\lambda_2 x - 1)\Bigg|_0^\infty = -\frac{1}{\lambda_2}.$$

Er volgt dus $\lambda_2 = -1/\lambda$ en daarmee wordt $p(x)$:

$$p(x) = -\frac{1}{\lambda}\exp\left(\frac{-x}{\lambda}\right),$$

hetgeen de negatief-exponentiële verdeling is.

**5.7.** *a.*

$$H(X) = -\int_{-\infty}^{\infty} \frac{1}{2\lambda}\exp\left(-\frac{|x|}{\lambda}\right)\log\left[\frac{1}{2\lambda}\exp\left(-\frac{|x|}{\lambda}\right)\right]dx =$$

$$= -\int_{\infty}^{\infty} \frac{1}{2\lambda}\exp\left(-\frac{|x|}{\lambda}\right)\log\frac{1}{2} - \int_{-\infty}^{\infty} \frac{1}{2\lambda}\exp\left(-\frac{|x|}{\lambda}\right)\log\left[\frac{1}{\lambda}\exp\left(-\frac{|x|}{\lambda}\right)\right]dx =$$

$$= 1 - \int_0^\infty \frac{1}{\lambda}\exp\left(\frac{-x}{\lambda}\right)\log\left[\frac{1}{\lambda}\exp\left(\frac{-x}{\lambda}\right)\right]dx.$$

Beschouwen we opgave 5.6.a dan kan geconcludeerd worden dat de term met de integraal precies de hoeveelheid informatie van de negatief-exponentiële verdeling is. Daarmee geldt voor $H(X)$ in het onderhavige geval:

$$H(X) = 1 + \log\lambda + \log e.$$

*b.* Aangezien $\mathbf{y} = |\mathbf{x}|$ zal de kans op $y$ gelijk zijn aan de kans op $-x$ en $+x$. Aangezien $p(x) = p(-x)$ zal de kans op $\mathbf{y} = |\mathbf{x}|$ gelijk worden aan tweemaal die van $p(x)$. Met andere woorden:

$$q(y) = \frac{1}{\lambda}\exp\left(-\frac{y}{\lambda}\right), \qquad y \geq 0.$$

Vergelijken we dit resultaat met opgave 5.6.a dan volgt direct

$$H(Y) = \log \lambda + \log e.$$

*c*. Als **x** bekend is, is **y** bekend. Dat wil zeggen dat als **x** bekend is, is er geen onzekerheid over **y** en dus zal gelden $H(Y/X) = 0$.
*d.* Er geldt algemeen

$$H(X,Y) = H(X) + H(Y/X) = H(Y) + H(X/Y).$$

Substitueert men de eerder gevonden waarden voor $H(X)$, $H(Y)$ en $H(Y/X)$ dan volgt

$$H(X/Y) = 1.$$

Dat wil zeggen: als **y** bekend is, is er nog onzekerheid over **x**. Dit klopt, immers het teken is onbekend. **x** kan positief of negatief zijn.

**5.8.** *a.* Als $x \le \frac{1}{2}\sqrt{2}$ dan geldt **z** = **y**. Dus geldt:

$$H(Z/x < \tfrac{1}{2}\sqrt{2}) = H(Y/x < \tfrac{1}{2}\sqrt{2}) = H(Y).$$

Voor de kansdichtheid $p(y)$ geldt $p(y) = \sqrt{2} - y$ voor $0 \le y \le \sqrt{2}$. De hoeveelheid informatie kan nu als volgt berekend worden:

$$H(Y) = -\int_0^{\sqrt{2}} (\sqrt{2} - y)\cdot\log(\sqrt{2} - y)\,\mathrm{d}y =$$

$$= \int_{\sqrt{2}}^{0} t\cdot\log t\,\mathrm{d}t = \frac{1}{\ln 2}\, t^2\left(\tfrac{1}{2}\ln t - \tfrac{1}{4}\right)\Big|_{\sqrt{2}}^{0} =$$

$$= -{}^2\log\sqrt{2} + \frac{1}{2\ln 2} = \tfrac{1}{2}\log\frac{e}{2} \text{ bit.}$$

*b.* Er geldt aangezien **z** = **x** voor $\mathbf{x} \ge \frac{1}{2}\sqrt{2}$ dat

$$H(Z/x \ge \tfrac{1}{2}\sqrt{2}) = H(X/x \ge \tfrac{1}{2}\sqrt{2}).$$

Met de formule van Bayes volgt:

$$p(x/x \ge \tfrac{1}{2}\sqrt{2}) = \frac{p(x)\cdot p(x \ge \frac{1}{2}\sqrt{2}/x)}{p(x \ge \frac{1}{2}\sqrt{2})} =$$

$$= \frac{p(x)}{p(x \ge \frac{1}{2}\sqrt{2})} = \tfrac{1}{2}\,p(x) \quad \text{voor} \quad \tfrac{1}{2}\sqrt{2} < x < \sqrt{2}.$$

Hieruit volgt, tezamen met $p(x) = x$, dat

$$H(X/x \geq \tfrac{1}{2}\sqrt{2}) = -\int_{\frac{1}{2}\sqrt{2}}^{\sqrt{2}} \tfrac{4}{3}\, p(x)\cdot\log\left[\tfrac{4}{3}\, p(x)\right] \mathrm{d}x =$$

$$= -\int_{\frac{1}{2}\sqrt{2}}^{\sqrt{2}} \tfrac{4}{3}\, x\cdot\log\left[\tfrac{4}{3}\, x\right] \mathrm{d}x = -\int_{\frac{2}{3}\sqrt{2}}^{\frac{4}{3}\sqrt{2}} \tfrac{3}{4}\, t\cdot\log t\, \mathrm{d}t = -\frac{3}{4\ln 2}\, \mathrm{t}^2\left(\tfrac{1}{2}\ln t - \tfrac{1}{4}\right)\Big|_{\frac{2}{3}\sqrt{2}}^{\frac{4}{3}\sqrt{2}} =$$

$$= -\tfrac{8}{3}\left[\tfrac{1}{2}\log(4\sqrt{2}) - \tfrac{1}{2}\log 3 - \tfrac{1}{4}\log e\right] + \tfrac{2}{3}\left[\tfrac{1}{2}\log(2\sqrt{2}) - \tfrac{1}{2}\log 3 - \tfrac{1}{4}\log e\right]$$

$$= -\frac{16}{7} + \log(3\sqrt{e})\ \text{bit}.$$

*c.* $H(Z) = p(x < \tfrac{1}{2}\sqrt{2})\cdot H(Z/x < \tfrac{1}{2}\sqrt{2}) + p(x \geq \tfrac{1}{2}\sqrt{2})\cdot H(Z/x > \tfrac{1}{2}\sqrt{2})$.
Aangezien $p(x < \tfrac{1}{2}\sqrt{2}) = \tfrac{1}{4}$ en $p(x \geq \tfrac{1}{2}\sqrt{2}) = \tfrac{3}{4}$ volgt tezamen met de eerder gevonden waarden in a. en b. dat:

$$H(Z) = \tfrac{1}{4}\cdot\tfrac{1}{2}\log\frac{e}{2} + \tfrac{3}{4}\cdot\left[-\frac{17}{6} + \log(3\sqrt{e})\right] = -\tfrac{1}{2}\log e - \tfrac{9}{4} + \tfrac{3}{4}\log 3.$$

**5.9.** *a.* Er geldt voor $p(x,y)$:

$$p(x,y) = \frac{1}{32} \quad \text{voor} \quad 0 \leq x \leq 8 \ \text{en}\ 0 \leq y \leq 8-x.$$

De kansdichtheid $p(x)$ kan als volgt berekend worden:

$$p(x) = \int_0^{8-x} p(x,y)\,\mathrm{d}y = \frac{8-x}{32}.$$

De hoeveelheid informatie $H(X)$ is dan gelijk aan:

$$H(X) = -\int_0^8 \frac{8-x}{32}\log\left(\frac{8-x}{32}\right)\mathrm{d}x = \int_{1/4}^{0} 32\cdot t\cdot\log t\,\mathrm{d}t =$$

$$= 32\,\frac{t^2}{\ln 2}\left(\tfrac{1}{2}\ln t - \tfrac{1}{4}\right)\Big|_{1/4}^{0} = 2 + \tfrac{1}{2}\log e.$$

b. Er geldt $p(y/x) = p(x,,y)/p(x) = 1/(8–x)$, waarbij $0 < y < 8-x$. Voor de conditionele hoeveelheid informatie volgt nu:

$$H(Y/X) = -\int_0^8\int_0^{8-x} p(x,y)\cdot\log[p(y/x)]\,\mathrm{d}y\mathrm{d}x =$$

$$= \int_0^8\int_0^{8-x} \frac{1}{32}\log(8-x)\,\mathrm{d}y\mathrm{d}x = \frac{1}{32}\int_0^8 (8-x)\cdot\log(8-x)\,\mathrm{d}x =$$

$$= -\frac{1}{32}\int_8^0 t\cdot\log t \, dt = -\frac{1}{32}\frac{t^2}{\ln 2}\left(\frac{1}{2}\ln t - \frac{1}{4}\right)\Big|_8^0 = 3 - \frac{1}{2}\log e.$$

*c*. Met behulp van de antwoorden van a. en b. volgt:

$$H(X,Y) = H(X) + H(Y/X) = 2 + \frac{1}{2}\log e + 3 - \frac{1}{2}\log e = 5 \text{ bit.}$$

Directe berekening levert

$$H(X,Y) = \int_0^8 \int_0^{8-x} p(x,y)\cdot\log p(x,y)\, dydx = \int_0^8 \int_0^{8-x} \frac{1}{32}\log\frac{1}{32}\, dxdy = 5 \text{ bit.}$$

**5.10.** *a*. Aangezien $p(x,y)$ de twee-dimensionale gaussische kansdichtheid is, waarbij $E(\mathbf{y}) = 0$ en $\mathrm{var}(\mathbf{y}) = \sigma_2^2$, volgt voor $q(y)$:

$$q(y) = \frac{1}{\sigma_2\sqrt{2\pi}}\cdot \exp(-\frac{y^2}{2\sigma_2^2}).$$

Op grond van stelling 5.2 volgt

$$H(Y) = \log(\sigma_2\sqrt{2\pi e}).$$

*b*. De kansdichtheid $p(x/y)$ laat zich herleiden uit

$$p(x/y) = p(x,y)/q(y) =$$

$$= \frac{1}{\sigma_1\sqrt{2\pi}\sqrt{1-\rho^2}}\cdot\exp\left[-\frac{1}{2(1-\rho^2)}\left\{\frac{x^2}{\sigma_1^2} - \frac{2\rho xy}{\sigma_1\sigma_2} + \frac{y^2}{\sigma_2^2}\right\} + \frac{y^2}{2\sigma_2^2}\right] =$$

$$= \frac{1}{\sqrt{2\pi}\sigma_1\sqrt{1-\rho^2}}\cdot\exp\left[-\frac{(x-\frac{\sigma_1}{\sigma_2}\rho y)^2}{2\sigma_1^2(1-\rho^2)}\right].$$

Dit is weer een gaussische kansdichtheid met conditioneel gemiddelde $(\sigma_1/\sigma_2)\rho y$ en variantie $\sigma_1^2(1-\rho^2)$.

Voor $H(X/Y)$ volgt nu:

$$\int_{-\infty}^{\infty} q(y)\left\{\int_{-\infty}^{\infty} p(x/y)\log p(x/y)\, dx\right\} dy,$$

waarbij $p(x/y)$ gelijk is aan boven gegeven gaussische kansdichtheid. Aangezien, zoals eenvoudig aangetoond kan worden, de hoeveelheid informatie in geval van een gaussische kansdichtheid alleen afhangt van de variantie en niet van het gemiddelde, volgt met behulp van stelling 5.2 dat

$$\int_{-\infty}^{\infty} p(x/y) \log p(x/y)\, dy = \log \left[ \sigma_1 \sqrt{1-\rho^2}\, \sqrt{2\pi e} \right]$$

en dus geldt ook:

$$H(X/Y) = \log \left[ \sigma_1 \sqrt{1-\rho^2}\, \sqrt{2\pi e} \right]$$

*c.* Aangezien $H(X,Y) = H(Y) + H(X/Y)$ volgt met de resultaten van a. en b.:

$$H(X,Y) = \log\left[\sigma_2 \sqrt{2\pi e}\right] + \log\left[\sigma_1 \sqrt{1-\rho^2}\, \sqrt{2\pi e}\right] =$$

$$= \log\left[\sigma_1 \sigma_2 \sqrt{1-\rho^2}\, 2\pi e\right].$$

*d.* Als $\rho = 0$ dan betekent dit dat ($\mathbf{x}$ – E($\mathbf{x}$)) en ($\mathbf{y}$) – E($\mathbf{y}$)) geen samenhang vertonen. Aangezien gegeven is dat E($\mathbf{x}$) = E($\mathbf{y}$) = 0 volgt dus dat $\mathbf{x}$ en $\mathbf{y}$ geen samenhang vertonen. Met andere woorden: $\mathbf{x}$ en $\mathbf{y}$ zijn onafhankelijk van elkaar.
Substitutie van $\rho = 0$ in $H(X/Y)$ levert:

$$H(X/Y) = \log\left[\sigma_1 \sqrt{2\pi e}\right],$$

hetgeen identiek is met $H(X)$. Dit stemt overeen met het feit dat voor $\rho = 0$ $\mathbf{x}$ en $\mathbf{y}$ onafhankelijk zijn.
Substitutie van $\rho = 0$ in $H(X,Y)$, zoals berekend in c., geeft:

$$H(X,Y) = \log\left[\sigma_1 \sigma_2\, 2\pi e\right] = H(X) + H(Y),$$

hetgeen wederom overeenstemt met de door $\rho = 0$ geïmpliceerde onafhankelijkheid van $\mathbf{x}$ en $\mathbf{y}$.