

Makers op waarde geschat

Een nieuwe selectiemethode voor de Nederlandse film

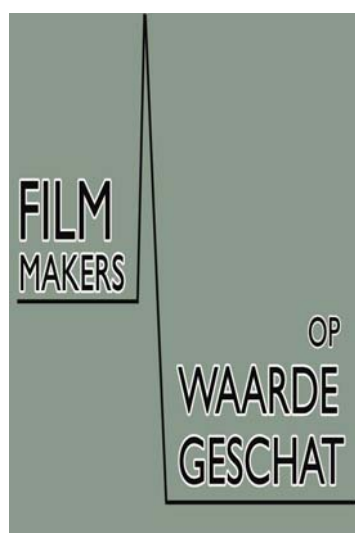
6 juni 2008

Auteurs:

Niek Baër
Maarten Bos
Arjan Feenstra

Begeleiders:

Esmé Lammers (Nederlands Fonds voor de Film)
Judith Vink-Timmer (Universiteit Twente)



Universiteit Twente
de ondernemende universiteit

Inhoudsopgave

1	Probleemstelling	3
1.1	Model anno september 2007	3
1.2	Modelvorming	3
2	Het model	5
2.1	Inleiding	5
2.2	Het model in woorden	6
2.2.1	Het cijfer van een film	6
2.2.2	Het cijfer van een filmmaker	7
2.2.3	Het cijfer van een samenwerking	8
2.2.4	Het cijfer van een filmteam	9
2.3	Het model in formules	9
2.3.1	Het cijfer van een film	9
2.3.2	Het cijfer van een filmmaker	10
2.3.3	Het cijfer van een samenwerking	14
2.3.4	Het cijfer van een filmteam	15
3	Analyse	16
3.1	Het bepalen van γ	16
3.2	Het bepalen van σ	17
4	Discussie	19
5	Aanbevelingen	20
6	Dankwoord	21
7	Bijlage	23

Inleiding

De Nederlandse taal is voor 23 miljoen mensen de moedertaal en voor nog eens vier miljoen mensen een tweede taal. Het Engels daarentegen wordt gesproken door 380 miljoen mensen als moedertaal en 130 miljoen als tweede taal [1]. Dit betekent dat er voor een Engelstalige film ruim 18 keer zoveel potentiële kijkers zijn. Zo leverde de film “Pirates of the Caribbean: At World’s End” in het openingsweekend in de Verenigde Staten meer dan 139 miljoen dollar op. De film “Alles is liefde” zet daar een omzet tegenover van 621.225,- euro in het openingsweekend, [2]. Hieruit blijkt goed dat het voor Nederlandse filmmakers veel lastiger is om geld te verdienen met het maken van films. Waarschijnlijk zouden er zonder steun vanuit de overheid alleen zeer goedkope producties gemaakt worden die gegarandeerd een groot publiek trekken, zoals “Kabouter Plop”. Om de Nederlandse film niet te laten verzanden in de massa van Amerikaanse en andere buitenlandse films, hebben Nederlandse filmmakers dus een steuntje in de rug nodig.

Om Nederlandse filmmakers aan te moedigen om commerciële en artistieke kwaliteitsfilms te maken, is er een fonds welke subsidies toekent aan de makers van een film. Deze doet dat via een aantal regelingen. Eén van die regelingen is de uitvoeringsregeling. Voor deze regeling, samen met de suppletierregeling, is er jaarlijks 40 miljoen euro beschikbaar. De suppletierregeling is er om bij lange speelfilms waarvan al een groot deel gefinancierd is, de rest van de financiering aan te vullen (suppleren). De uitvoeringsregeling is er juist om films (lange speelfilms, maar ook animaties, korte films en documentaires) een subsidie te geven aan het begin van het filmproductieproces zodat de film überhaupt gemaakt kan worden.

In dit verslag zullen we een nieuw model voor het behandelen van subsidieaanvragen behandelen. Voor het model maken wij gebruik van data over alle films van het Nederlands Fonds van de Film. Deze data is vertrouwelijk en daarom niet opgenomen in dit verslag. Indien men deze data toch wil inzien dient men contact op moeten nemen met het Nederlands Fonds voor de Film.

Hoofdstuk 1

Probleemstelling

1.1 Model anno september 2007

Op dit moment kan een producent op twee manieren aan extra steun komen voor een nieuwe film. Als de financiering voor minstens 65% rond is, dan zal de overheid, zonder kwaliteitseisen, de rest van het geld betalen. Dit is alleen bedoeld voor lange speelfilms. Wanneer dit echter niet het geval is, zal de producent moeten aankloppen bij de uitvoeringsregeling. Het grote verschil met de suppletieregeling, naast dat het geen lange speelfilm hoeft te zijn en het al gefinancierde deel niet van belang is, is dat er hier wel naar kwaliteit wordt gekeken. Deze kwaliteit wordt vastgesteld door een commissie. De commissie bestaat uit mensen met veel filmkennis zoals andere producenten, regisseurs en scenaristen. De beoordeling vindt plaats aan de hand van het filmplan. Dit lijkt bij een kundige commissie een goede manier om vast te stellen of de nieuwe productie wel of geen subsidie zal krijgen.

1.2 Modelvorming

Op het eerste gezicht lijkt het dus dat de subsidie op de goede plek zal komen. Het probleem zit hem echter in het woord kwaliteit. Heel vaak is dit niet meer dan een mening en slechts zelden een onvoldongen feit. Wat de één een goede film vindt, is voor de ander een productie van een bedenkelijk niveau. Een goede definitie van kwaliteit is dan ook niet te geven wat er toe leidt dat per commissie de criteria voor subsidies verschillen. Daaruit volgt dat producenten niet meer weten waar het filmplan nu aan moet voldoen.

Verder wordt er bij dit systeem aangenomen dat een commissie aan de hand van een filmplan kan beoordelen hoe goed een film gaat worden en misschien nog belangrijker: hoe goed het publiek de film gaat vinden. Naar verwachting is dit onjuist. Een scenario is aan het begin van een productie nog maar een schim van de uiteindelijke film. Het is zeer onwaarschijnlijk dat aan de hand van alleen het scenario in een vroegtijdig stadium kan worden voorspeld wat voor een film het uiteindelijk gaat worden. Daarom krijgt de commissie, naast het scenario, extra informatie over de productie van de film. Hierdoor zou de commissie een beter oordeel kunnen vellen over de subsidieaanvraag. Producenten met meer ervaring voelen zich hierdoor echter aangetast in de vrijheid van hun productie.

Op dit moment maken ze films dus niet naar hun eigen normen, maar stemmen het filmplan af op de commissie zodat ze een grotere kans hebben op subsidietoewijzing. In de tussentijd heeft dit systeem geleid tot vooral discussies over kwaliteit en niet tot aantoonbaar betere films [3].

De vraag is dus hoe op een betere en objectievere manier een beslissing kan worden genomen over de verdeling van het beschikbare geld onder de subsidieaanvragen.

Hoofdstuk 2

Het model

2.1 Inleiding

Hoe kunnen we op een objectieve manier bekijken of een film het goed gaat doen of niet? Om hier inzicht in te krijgen gaan we eerst kijken wat allemaal van invloed is op het succes van een film. Daarna kunnen we kijken wat we daarvan gaan gebruiken in ons model. Zo zou je bijvoorbeeld kunnen denken aan promotie van de film. Hoe meer reclame er gemaakt wordt, hoe meer mensen er wel naar die film zullen gaan. Zo kan een gemiddelde film met 20% meer promotie ongeveer 10% meer bezoekers trekken [4]. Ook is aangetoond dat een aantal grote namen van acteurs en actrices een positief effect heeft op het aantal bezoekers van een film [5]. Wat misschien ook invloed heeft op het succes van een film is wat voor recensies de film krijgt. Hier zijn de studies over verdeeld. Zo blijkt er uit de één een grote correlatie tussen recensies en het aantal bezoekers [6], terwijl de ander helemaal geen uitwerking ervan ziet [7].

Het probleem is dat we van tevoren willen schatten of de aanvraag een goede film gaat worden. Op dat moment zijn er nog helemaal geen recensies beschikbaar. Ook de acteurs die gaan spelen zijn vaak nog niet bekend en datzelfde geldt voor het budget van de films. Dat kunnen we dus niet gebruiken in ons model. Wat we wel kunnen gebruiken is data van films uit het verleden van diezelfde filmmakers om te bepalen hoe goed deze film gaat worden. We gaan er vanuit dat een film zo goed gaat worden als de makers zijn.

Een nieuw ingediend filmplan wordt op vier aspecten bekeken worden door het Film Fonds:

- Staat van dienst van het filmteam
- Financiële interesse van een film distributeur
- Inhoudelijke beoordeling van het filmplan door deskundigen
- Beleidsplan

De staat van dienst is een maat voor hoe goed een filmteam is. Hier zullen ervaren filmmakers met een aantal goed scorende films op hun naam veel punten scoren. Hierdoor kunnen makers die zichzelf al bewezen hebben meer continuïteit geboden worden. De financiële interesse uit de markt zien we als een

goede graadmeter voor het publiek. Als een distributeur bereid is de film gedeeltelijk te financieren moet hij er vertrouwen in hebben dat de productie een succes zal worden. Naast een filmplan alleen op objectieve cijfers van de eerste twee aspecten te beoordelen, zal er nog steeds een commissie aan te pas komen die het scenario nader gaat bestuderen. Het filmplan is nu alleen veel minder doorslaggevend dan in de oude situatie omdat er op meer onderdelen punten behaald kunnen worden. Als laatste kan er een soort bonus verdiend worden als de film past bij het beleidsplan. Dit is zo, als het fonds een bepaald genre wil stimuleren.

Voor elk onderdeel zijn er punten te verdienen en als een filmplan boven een bepaalde grens komt zal er subsidie worden verlenen. In dit verslag presenteren wij ons model hoe wij de staat van dienst van een filmteam bepalen.

2.2 Het model in woorden

In de probleemstelling zijn we tot de conclusie gekomen dat we op zoek moeten naar een model waarmee de overheidssubsidie op een verantwoorde manier verdeeld kan worden over de aanvragen. In deze paragraaf zullen we het door ons gevonden model uitleggen aan de hand van omschrijvingen. Wij zullen het gebruik van formules daarom zoveel mogelijk vermijden. De lezers die geïnteresseerd zijn in de wiskunde achter ons model verwijzen we naar de volgende paragraaf waarin we het model wiskundig zullen onderbouwen en nader zullen toelichten met formules.

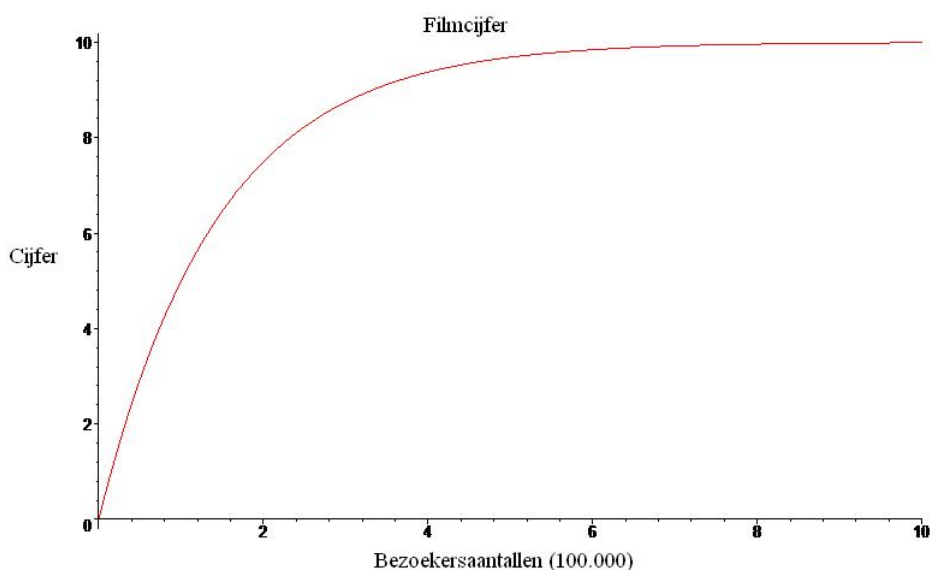
2.2.1 Het cijfer van een film

Wat we willen bereiken is een cijfer voor een compleet filmteam. In eerste instantie zit er in een filmteam altijd een producent, regisseur en scenarist. Dit zou dus betekenen dat zodra we een cijfer willen bepalen voor het filmteam we eerst een cijfer moeten bepalen voor de producent, regisseur en scenarist, oftewel voor elke filmmaker die in het filmteam zit. Om nu een cijfer te kunnen geven aan een filmmaker gaan we kijken wat we van deze filmmaker kunnen zeggen, het enige wat van een filmmaker bekend is, is dat hij films gemaakt heeft die al dan niet succesvol waren. De sleutel tot het vinden van een cijfer voor een filmmaker ligt dus in het vinden van een cijfer voor een film. Het cijfer voor een film moet aan de volgende eisen voldoen:

1. Het cijfer van een film moet tussen de 0 en 10 liggen. Dit heeft als resultaat dat een cijfer makkelijk te interpreteren is.
2. Wanneer een film 100.000 bezoekers of meer heeft beschouwen we deze film als redelijk succesvol en waarderen we een film met 100.000 bezoekers met een 5.
3. Wanneer een film alleen maar een prijs gewonnen heeft en er geen bezoekers zijn geweest beschouwen we deze film toch als redelijk succesvol. Een film met twee punten aan prijzen waarderen we daarom dus ook met een 5. Dit heeft als gevolg dat een prijs van twee punten gelijk staat aan 100.000 bezoekers.

4. Vanuit de filmwereld weten we dat het moeilijker is om van 50.000 bezoekers naar 100.000 bezoekers te gaan dan van 500.000 bezoekers naar 1.000.000 bezoekers. Daarom willen wij dit “vroeg” succes beter waarderen.

Wat we nu bepalen is een formule bepalen die al deze eisen in zich heeft. Het resultaat is hieronder te vinden. Wat we zien is dat deze formule snel stijgt bij een laag bezoekersaantal en zeer langzaam stijgt wanneer de film toch al een groot succes is. Precies wat wij willen volgens eis 4. Ook zien we dat het halen van 100.000 bezoekers (of een prijs van twee punten) wordt gewaardeerd met een 5, wat volgens eis 2 en 3 de bedoeling was. Als laatste is ook eis 1 makkelijk te controleren omdat het cijfer inderdaad tussen 0 en 10 komt te liggen volgens deze formule.



Figuur 2.1: Waarderingsformule

Wat we nu hebben is een formule die afhankelijk is van bezoekersaantallen en gewonnen prijzen. Hierbij moet er echter rekening gehouden worden met het feit dat bijv. een Oscar meer punten waard is dan een Gouden Kalf. Om toch te kunnen rekenen met deze prijzen werken we alles terug naar punten.

2.2.2 Het cijfer van een filmmaker

Van een filmmaker weten we zijn verleden. Welke films hij gemaakt heeft, wanneer ze geproduceerd zijn en welk cijfer ze gekregen hebben. Al die cijfers van zijn films gaan we combineren tot een verwachting van de films van die filmmaker. Die verwachting is het gemiddelde van zijn cijfers, maar niet elk cijfer telt even zwaar mee. We willen cijfers verder uit het verleden minder hard laten meetellen dan cijfers uit het recente verleden. Ook is het van belang of hij die film alleen heeft gemaakt of zijn gedeelte (productie, regie of scenario) samen met iemand anders heeft gedaan. Als we dit meenemen bij de cijfers meenemen, komen we op de verwachting van een filmmaker uit.

Naast een verwachting willen we ook graag een variantie weten zodat we de kans op succes kunnen berekenen. Dit is een soort zekerheid voor de filmmaker, hoe hoger de variantie, hoe minder zekerheid we in zijn cijfers hebben en hoe minder stabiel ze zijn. De variantie hangt omgekeerd evenredig samen met de ervaring van de filmmaker. Ervaring zien wij als de hoeveelheid films die een filmmaker heeft gemaakt met daarin meegenomen welk aandeel hij in deze film had en hoe lang geleden die zijn gemaakt. Hoe meer ervaring een filmmaker heeft, des te kleiner is zijn variantie. Uit de data kunnen we van iedere producent, regisseur en schrijver zijn ervaring halen. Daarmee weten we ook die variantie van alle filmmakers apart.

We weten van elke filmmaker zijn verwachting en zijn variantie. Nu willen we daar één cijfer van maken, die we de kans op succes noemen. Dit is de kans dat een filmmaker een nieuwe film gaat maken die een 5 of hoger scoort. Dat wil zeggen minstens 100.000 bezoekers of twee punten aan prijzen haalt. Omdat een kans altijd een getal tussen 0 en 1 is, doen we deze kans maal 10 om er een cijfer tussen 0 en 10 van te maken. Dit is het cijfer dat de filmmaker krijgt.

2.2.3 Het cijfer van een samenwerking

We hebben nu een manier om de kans op succes van één filmmaker te bepalen. Het komt echter vaak voor dat niet één, maar twee of meer producenten een film maken. Dit kan ook gelden voor de regisseur of de scenarist. Nu is de vraag hoe we een dergelijke samenwerking gaan beoordelen. Wat is hun gezamenlijke kans op succes? We willen dat een samenwerking de volgende gevolgen heeft op hun kans op succes:

- Een ervaren en goede producent moet geen last ondervinden van samenwerken met een onervaren persoon. Op deze manier wordt samenwerking en dus ontwikkeling van nieuwe filmmakers niet belemmerd.
- Twee goede producenten kunnen elkaar aanvullen en zo hun kans op succes vergroten.

We beginnen met het bepalen van een verwachting voor de samenwerking van twee producenten. Hier kunnen we niet gewoon het gemiddelde nemen van de verwachtingen van de twee producenten. Dit zou flinke nadelige gevolgen hebben voor de ervaren producent als de onervaren producent een lage verwachting heeft. We omzeilen dit door de films van de twee producenten bij elkaar te nemen en daar op dezelfde manier als bij één producent de verwachting te bepalen. Dit heeft als voordeel dat de cijfers van de ervaren filmmaker zwaarder meetellen dan die van de onervaren. Daarnaast is er nog een belangrijke factor: de beschikbaarheid. Deze geeft aan hoeveel tijd een filmmaker in de film wil steken. De films van een producent die een hogere beschikbaarheid heeft zullen in de verwachting zwaarder meetellen.

Naast de verwachting hebben we ook de variantie, oftewel zekerheid, van een samenwerking nodig. Deze bepalen we op een soortgelijke manier als de verwachting. We kunnen nu de variantie van de twee bij elkaar genomen filmmakers bepalen zoals we dat ook bij één filmmaker deden. De gezamenlijke ervaring is dan minstens even groot als die van de meest ervaren producent. De variantie van een samenwerking is dus altijd kleiner of gelijk aan de variantie

van elke persoon apart. Hier komt het aanvullen van twee producenten naar boven. Door de grotere ervaring zal de kans op succes toenemen.

2.2.4 Het cijfer van een filmteam

De cijfers voor elk onderdeel van het filmteam (productie, regie en scenario) zijn nu bekend. Het cijfer van het hele filmteam samen is te berekenen. Via het filmfonds weten we dat niet alle cijfers even zwaar meewegen. Een producent moet zwaarder wegen dan een regisseur, die op zijn beurt weer meer aandeel heeft dan een scenarist. Dit doen we door middel van een weging van respectievelijk drie staat tot twee staat tot één. Het cijfer voor een filmteam is dan de som van drie keer het cijfer van de producenten, twee keer het cijfer van de regisseurs en één keer het cijfer van de scenaristen. Dit levert een puntenaantal tussen de 0 en 60 op wat het puntenaantal van de staat van dienst is.

2.3 Het model in formules

In deze paragraaf zullen we het door ons gevonden model onderbouwen en uitleggen aan de hand van formules. De lezers die niet geïnteresseerd zijn in de wiskunde achter ons model kunnen deze paragraaf overslaan zonder de rode draad van het model te verliezen.

2.3.1 Het cijfer van een film

Voordat we een filmmaker een cijfer kunnen geven moeten we eerst weten hoe goed zijn films waren. Om dit te doen moet er een waarderingsformule bepaald worden die aan de volgende eisen voldoet:

1. Het cijfer van een film moet tussen de 0 en 10 liggen.
2. Wanneer een film 100.000 bezoekers of meer heeft beschouwen we deze film als redelijk succesvol en waarderen we deze met minstens een 5.
3. Een film moet een 5 opleveren bij het winnen van een prijs van 2 punten.
4. Het verschil tussen 50.000 en 100.000 bezoekers moet beter gewaardeerd worden dan het verschil tussen 500.000 en 1.000.000 bezoekers.

Wanneer we naar eis 4 en naar eis 1 kijken dan beginnen we al snel aan een e-macht te denken in de volgende vorm:

$$10(1 - e^{-x})$$

Als we echter willen voldoen aan eis 2 en eis 3 zien we dat onze aanname van een e-macht niet de meest optimale keuze was. We introduceren twee variabelen B en P , met B het aantal bezoekers in 100.000-tallen en P het aantal punten verkregen door het winnen van prijzen. Dan begint onze formule er uit te zien als:

$$10(1 - x^{-B-P})$$

Wanneer we nu kijken naar eis 2 en eis 3 dan zien we al snel in dat een prijs van twee punten gelijk staat aan 100.000 bezoekers. Onze gekozen P moeten we dus veranderen in $\frac{1}{2}P$.

$$10(1 - x^{-B - \frac{1}{2}P})$$

Wat we nu weten is dat een film met 100.000 bezoekers, $B = 1$, gewaardeerd wordt met een 5. Daardoor kunnen we nu x bepalen:

$$\begin{aligned} 10(1 - x^{-1}) &= 5 \\ x &= 2 \end{aligned}$$

We hebben nu dus de volgende formule gevonden voor het waarderen van een film:

$$C_j = 10 \left(1 - 2^{-B_j - \frac{1}{2}P_j} \right) \quad (2.1)$$

met:

- C_j het cijfer van film j
- B_j het aantal bezoekers van film j in 100.000-tallen
- P_j het aantal met prijzen behaalde punten van van film j

Er moet echter wel op gelet worden dat prijzen onderling in waarde verschillen, om hiermee te kunnen rekenen moeten de prijzen omgezet worden naar punten.

2.3.2 Het cijfer van een filmmaker

Op dit moment hebben we nu voor elke film die gemaakt is door een filmmaker een cijfer. Deze cijfers combineren we om zo voor elke filmmaker apart een cijfer te bepalen. Om dit te kunnen doen gaan we er vanuit dat het cijfer van elke filmmaker een zekere kansverdeling heeft. Aan de hand van deze kansverdeling bepalen we een kans dat een filmmaker voor zijn volgende film een 5 of hoger scoort. Voor we deze kans kunnen berekenen moet we weten hoe het kansverdeling eruit ziet.

Het bepalen van een kansmodel

Voor het schatten van de kans op succes hebben we een verwachting en een variantie van een filmmaker nodig. Hier stuiten we echter op een probleem. Elke filmmaker heeft apart een eigen variantie maar het is onmogelijk om die op een goede manier te verkrijgen. Dit komt doordat elke filmmaker apart hoogstens een stuk of 10 films gemaakt heeft en in veel gevallen een stuk minder. Dit is echter te weinig data om een gegronde schatting te maken. Om toch tot een goede schatting van de variantie te komen gaan we gebruik maken van een additief model. Dat wil zeggen dat elke filmmaker een eigen verwachting heeft, maar filmmakers hebben wel een gezamenlijke variantie. Hierop zullen echter wat afwijkingen optreden ten gevolge van de tijd. Zo willen we aan een wat oudere film minder zwaar tillen dan aan een nieuwere. Dit kan uitgelegd worden als een trekking uit een verdeling waarvan de variantie groter is. Ook heeft de ervaring invloed op de variantie. Zo zal naar verwachting een ervaren filmmaker constanter presteren dan een beginneling.

In dit model is X_{ij} de j^e film van filmmaker i . De verwachting van filmmaker i noemen we μ_i en U_{ij} is de uitwijking van X_{ij} ten opzichte van μ_i . Als we dan μ_i gaan schatten met $\widehat{\mu}_i$, kunnen we de uitwijking U_{ij} schatten met \widehat{U}_{ij} .

$$\begin{aligned} X_{ij} &= \mu_i + U_{ij} \\ X_{ij} &\approx \widehat{\mu}_i + \widehat{U}_{ij} \end{aligned}$$

We weten nu dat de uitwijking U_{ij} een verwachtingswaarde gelijk aan 0 heeft en een tot nu toe onbekende variantie. Van deze variantie weten we dat deze afhankelijk van de ervaring van filmmaker i , in de vorm van w_i . Daarnaast is deze variantie afhankelijk van een tijdsvariabele v_j voor elke film j omdat er in ons model rekening wordt gehouden met een verdiscontering. Verder nemen we ook aan dat elke filmmaker een vaste variantie σ^2 heeft. Dit leidt samen tot het volgende model:

$$E(U_{ij}) = 0$$

$$Var(X_{ij}) = Var(\mu_i + U_{ij}) = Var(U_{ij}) = \sigma^2 w_i^2 v_j^2$$

Vervolgens corrigeren we elke uitwijking over deze ervaring w_i en tijdsvariabele v_j zodat we de gegevens van alle filmmakers kunnen gebruiken om σ^2 te bepalen:

$$\begin{aligned} \frac{X_{ij}}{v_j w_i} &= \frac{\mu_i}{v_j w_i} + \frac{U_{ij}}{v_j w_i} \\ \frac{X_{ij} - \mu_i}{v_j w_i} &= \frac{U_{ij}}{v_j w_i} = R_{ij} \\ \widehat{R}_{ij} &= \frac{X_{ij} - \widehat{\mu}_i}{v_j w_i} \end{aligned}$$

Nu is R_{ij} gestandaardiseerd met verwachting 0 en variantie σ^2 . Met behulp van Maple kunnen we nu de variantie van R_{ij} schatten. Vervolgens kunnen we ook uit het histogram van R_{ij} aflezen wat voor verdeling de data zou hebben. Voordat we σ^2 echter kunnen bepalen moeten we nog drie variabelen bepalen, namelijk v_j , $\widehat{\mu}_i$, en w_i .

Het bepalen van v_j , $\widehat{\mu}_i$ en w_i

In het kansmodel is een tijdsvariabele v_j^2 gedefinieerd die we gebruiken om de variantie van een filmmaker te bepalen. Deze v_j^2 is specifiek voor elke film en geeft een bepaalde zekerheid aan van het behaalde filmcijfer. Zodra er een film is die erg lang geleden is gemaakt, dan kunnen we weinig zekerheid geven over het kunnen van de filmmaker. We kunnen dit modelleren met een verdisconteringswaarde, elk jaar hebben we minder zekerheid over het cijfer. Ook kunnen we weinig zekerheid geven wanneer de filmmaker voor maar een klein aandeel in de film heeft gewerkt. Een goede manier om dit aandeel te weten te komen is via het winstaandeel. Laten we nu het winstaandeel van een filmmaker in film j α_j noemen en de verdisconteringswaarde γ . Als we nu gaan kijken naar het effect van α en γ op de zekerheid dan vinden we een omgekeerd evenredig verband, welke aanleiding zou geven tot de volgende formule voor v_j^2 :

$$v_j^2 = \frac{1}{\alpha_j \gamma^{t-t_i}} \quad (2.2)$$

Vervolgens willen we een convexe combinatie vinden tussen alle films van een filmmaker waarvan de variantie zo klein mogelijk is. We zijn dus op zoek naar de beste zuivere schatter voor μ_i met deze v_j . Om deze zuivere schatter te vinden zijn we op zoek naar de oplossing van het volgende probleem waarin we een som nemen over alle films j van filmmaker i :

$$\min \left\{ \sum_{j \in f_i} d_j^2 v_j^2 \mid \sum_{j \in f_i} d_j = 1 \right\}$$

Waarbij d_k de fractie weergeeft waarin cijfer C_k voorkomt in de (zuivere) schatter voor μ_i . Met behulp van Lagrange Multipliers kunnen we dit omzetten naar het probleem zonder restricties:

$$\min \left\{ \sum_{j \in f_i} d_j^2 v_j^2 + \lambda \left(\sum_{j \in f_i} d_j - 1 \right) \right\}$$

Welke we kunnen oplossen door de volgende partiële afgeleides gelijk te stellen aan 0 en op te lossen. Voor elke film j van filmmaker i , $j \in F_i$, krijgen we dan:

$$\frac{\partial}{\partial d_j} \left(\sum_{j \in f_i} d_j^2 v_j^2 + \lambda \left(\sum_{j \in f_i} d_j - 1 \right) \right) = 2d_j v_j^2 + \lambda = 0$$

oftewel:

$$d_j = \frac{-\lambda}{2v_j^2} \quad (2.3)$$

Voor de partiële afgeleide naar λ vinden we:

$$\frac{\partial}{\partial \lambda} \left(\sum_{j \in f_i} d_j^2 v_j^2 + \lambda \left(\sum_{j \in f_i} d_j - 1 \right) \right) = \sum_{j \in f_i} d_j - 1 = 0$$

oftewel:

$$\sum_{j \in f_i} d_j = 1 \quad (2.4)$$

Het combineren van (2.3) en (2.4) geeft:

$$\sum_{j \in f_i} \frac{-\lambda}{2v_j^2} = 1$$

oftewel:

$$\lambda = \frac{-2}{\sum_{j \in f_i} \frac{1}{v_j^2}}$$

Samen met (2.3) geeft dit:

$$\begin{aligned}
d_j &= \frac{-\lambda}{2v_j^2} \\
&= \frac{-\lambda}{\frac{2}{\alpha_j \gamma^{t-t_j}}} \\
&= \frac{2}{\sum_{k \in f_i} \alpha_k \gamma^{t-t_k}} \\
&= \frac{\alpha_j \gamma^{t-t_j}}{\sum_{k \in f_i} \alpha_k \gamma^{t-t_k}} \tag{2.5}
\end{aligned}$$

Nu bekend is wat d_j moet zijn kunnen we de beste zuivere schatter geven voor μ_i volgens:

$$\begin{aligned}
\hat{\mu}_i &= \sum_{j \in f_i} d_j C_j \\
&= \sum_{j \in f_i} \frac{\alpha_j \gamma^{t-t_j}}{\sum_{k \in f_i} \alpha_k \gamma^{t-t_k}} C_j \tag{2.6}
\end{aligned}$$

Wat neerkomt op een gewogen gemiddelde van alle cijfers. We hebben nu een expressie voor v_j en $\hat{\mu}_i$, wat nog mist is w_i . De variabele w_i geeft een maat van zekerheid aan welke nauw samen loopt met de ervaring van een filmmaker. Hoe meer ervaring filmmaker i heeft, hoe meer zekerheid wij kunnen geven over zijn cijfer. We zijn dus op zoek naar een expressie voor de ervaring van een filmmaker. Gegeven is dat filmmaker i een aantal films op zijn naam heeft staan en van elke film is bekend hoe lang deze geleden is en wat het winstaandeel van filmmaker i in deze film was. We hebben dus voor elke film de volgende waarde:

$$\alpha_j \gamma^{t-t_j}$$

Wanneer we deze waarden voor elke film uit het oeuvre van filmmaker i optellen dan hebben we een goede graadmeter voor de ervaring van filmmaker i . We kiezen dus:

$$w_i^2 = \frac{1}{\sum_{j \in f_i} \alpha_j \gamma^{t-t_j}} \tag{2.7}$$

met:

- f_i de verzameling films van filmmaker i

Het bepalen van een cijfer

Uit de vorige paragrafen is gebleken dat we een kansmodel hebben voor een filmmaker. In dit kansmodel hebben we gevonden:

$$\hat{\mu}_i = \sum_{j \in f_i} \frac{\alpha_j \gamma^{t-t_j}}{\sum_{k \in f_i} \alpha_k \gamma^{t-t_k}} C_j \tag{2.8}$$

Vervolgens moet er nog een variantie bepaald worden voor een specifieke filmmaker, bekend is dat de variantie gegeven wordt door:

$$Var(X_{ij}) = \sigma^2 w_i^2 v_j^2$$

Nu is v_j^2 echter van de vorm uit (2.2) en afhankelijk van de tijd door zijn verdiscontering en afhankelijk van de film door het winstaandeel en moment van uitgave. Om nu een v_j voor een filmmaker te bepalen, bepalen we v_{j_0} waarin we net doen alsof we een film bekijken uit het huidige jaartal en waarin de filmmaker alleen heeft gewerkt. In dat geval wordt deze v_{j_0} gelijk aan 1. We krijgen dus een variantie voor een filmmaker gelijk aan:

$$\text{Var}(V_i) = \sigma^2 w_i^2 = \frac{\sigma^2}{\sum_{j \in f_i} \alpha_j \gamma^{t-t_j}} \quad (2.9)$$

met:

- V_i de verdeling van filmmaker i .

Uit het histogram van de R_{ij} halen we de verdeling F_i voor elke filmmaker i , dit histogram is te vinden in de analyse. Met behulp van deze verdeling definiëren we:

$$C_{F_i} = 10 \cdot P(V_i \geq 5) \quad (2.10)$$

2.3.3 Het cijfer van een samenwerking

In de vorige paragraaf zijn we op een model uitgekomen om het cijfer voor één filmmaker te bepalen. Nu willen we een manier vinden om een samenwerking tussen een aantal filmmakers te waarderen. Deze samenwerking willen wij op dezelfde manier waarderen als een filmmaker die alleen werkt. Weer gaan we dus op zoek naar een verwachtingswaarde en een variantie. Voor de verwachtingswaarde van een samenwerking S tussen n filmmakers, $E(S)$, willen we graag alle cijfers meenemen maar ook een zekere beschikbaarheid β van een filmmaker. Elke filmmaker kan namelijk een aantal punten op een film inzetten met een maximum van 200 punten (procent) per jaar. Laten we B_i het aantal punten noemen wat filmmaker i beschikbaar stelt voor een nieuwe film. Om dat een β te bepalen kijken we naar elke B_i van elke filmmaker i in de samenwerking S . We definiëren β_k van filmmaker k als

$$\beta_k = \frac{B_k}{\sum_{i \in S} B_i}$$

Wat we nu willen doen is wederom een gewogen gemiddelde nemen over alle cijfers van de filmmakers met de gewichten α en γ maar dit keer ook met het gewicht β per filmmaker. We vinden nu een nieuwe formule om de verwachtingswaarde van een samenwerking S te bepalen:

$$E(V_S) = \frac{\sum_{i \in S} \beta_i \sum_{j \in f_i} \alpha_j \gamma^{t-t_j} C_j}{\sum_{i \in S} \beta_i \sum_{j \in f_i} \alpha_i \gamma^{t-t_j}}$$

met:

- V_S de verdeling van de samenwerking S
- β_i de beschikbaarheidsfactor van filmmaker i
- C_j het cijfer van film j

Op dezelfde manier vinden we ook een variantie. We willen hierbij alleen als eis stellen dat zodra er meerdere filmmakers samenwerken zij een kleinere variantie vinden dan wanneer zij allen alleen werken. Hierdoor vinden we:

$$Var(V_S) = \frac{\sigma^2}{\sum_{i \in S} \sum_{j \in f_i} \alpha_j \gamma^{t-t_j}}$$

2.3.4 Het cijfer van een filmteam

De cijfers voor elk onderdeel van het filmteam (productie, regie en scenario) zijn nu bekend. Het cijfer van het hele filmteam samen is zoals uitgelegd wordt in 2.2.4 op de volgende manier te berekenen:

$$\frac{3C_P + 2C_R + C_S}{6}$$

Met C_P het cijfer voor de producent (het producententeam), C_R het cijfer voor de regisseur (het regisseurenteam) en C_S het cijfer voor de scenarist (het scenaristenteam). Dit cijfer is wederom een cijfer tussen 0 en 10. In het huidige model willen wij 60 punten verdelen voor de staat van dienst. Het verkregen cijfer kunnen we dus maal 6 doen en we hebben het aantal punten wat het filmteam krijgt in de beoordeling.

Hoofdstuk 3

Analyse

In dit hoofdstuk zullen we de benodigde numerieke variabelen bepalen. In ons model zijn dit γ en σ .

3.1 Het bepalen van γ

De verwachting van een filmmaker willen we af laten hangen van de films die hij gemaakt heeft. Maar niet elke film zegt evenveel over zijn kunnen. Recentere films laten beter zien hoe goed hij of zij nu is dan films van 30 jaar oud. Om resultaten van oude films minder zwaar mee te laten tellen hebben we een verdisconteringswaarde γ . Deze variabele zorgt ervoor dat een film elk jaar een bepaalde factor minder waard wordt. Maar hoe bepalen we de waarde van γ ? Dat ligt eraan hoeveel waarde we nog aan een film hechten na zoveel jaar. Een mooie manier om dat te verwoorden is de halveringstijd. Wanneer hecht je nog maar half zoveel waarde aan een film ten opzichte van wanneer hij net uit was? Stel dat we willen dat een film na 10 jaar nog maar de helft waard moet zijn. Dat levert de volgende γ op.

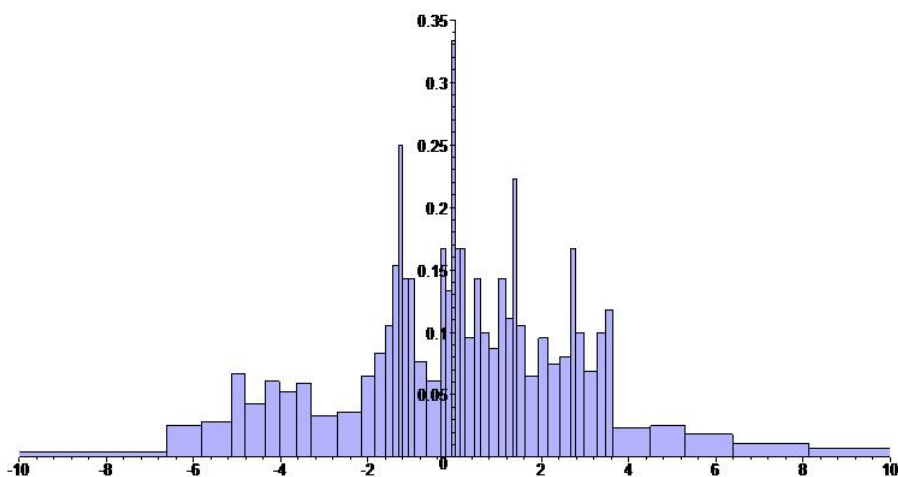
$$\begin{aligned}\gamma^{10} &= \frac{1}{2} \\ \gamma &= 0.933\end{aligned}$$

Algemeen, wanneer we een halveringstijd van T jaren willen hebben dan kiezen we:

$$\gamma = \sqrt[T]{\frac{1}{2}}$$

3.2 Het bepalen van σ

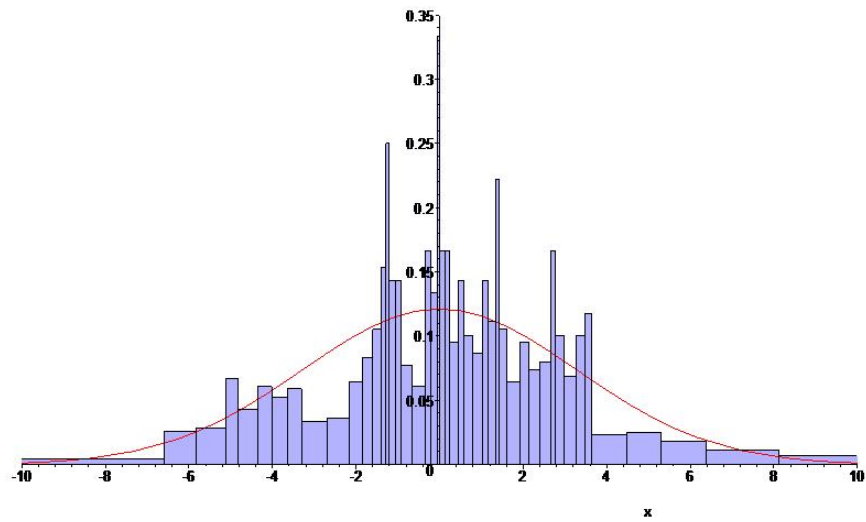
Zoals al eerder vermeld, gaan we alle films gebruiken om een gezamenlijke sigma te schatten. Hiertoe nemen we de cijfers van alle films en trekken hier de verwachting vanaf. Na de nodige correctie op de restterm U_{ij} te hebben uitgevoerd (deling door w_i en v_j) komen we tot de R_{ij} . Deze waarden plotten we allemaal in een histogram om te zien hoe deze verdeeld zijn. Onze schatter $\hat{\sigma}$ is de steekproefstandaardafwijking van deze dataset. Deze kunnen we met Maple uitrekenen en dat levert ons een $\hat{\sigma}$ van 3,296. Het histogram met daarin de waarden voor R_{ij} is hieronder te vinden:



Figuur 3.1: Histogram

Zodra we weten dat R_{ij} normaal verdeeld is kunnen we ook de conclusie trekken dat X_{ij} dat moet zijn. Uit het histogram hierboven blijkt dat een normale verdeling inderdaad de beste keus is. Er zitten wel een aantal uitschieters bij, maar we hebben eenvoudigweg te weinig data om de normale verdeling te verwerpen. We proberen dan ook de onbekende verdeling te modelleren met een normale verdeling.

Nu we $\hat{\sigma}$ hebben kunnen we de verdeling met verwachting 0 en variantie $\hat{\sigma}$ ook plotten. Als we deze over het histogram plotten krijgen we het volgende plaatje:



Figuur 3.2: Histogram met normale verdeling

We zien dat de normale verdeling niet helemaal de data benadert, maar er zou meer data moeten zijn om hier een gegronde uitspraak over te doen. De Maple-sheet waarin dit alles is uitgewerkt is bijgevoegd en is te vinden bij de bijlagen.

Hoofdstuk 4

Discussie

In dit hoofdstuk bespreken we onze belangrijkste resultaten. Deze resultaten zijn echter vertrouwelijk en daardoor niet opgenomen in dit verslag. Indien men deze resultaten toch in wil zien, dient met contact op te nemen met het Nederlands Fonds voor de Film.

Van tevoren was het doel om een nieuwe selectieprocedure te maken voor subsidieaanvragen. Daarin was het van belang dat filmmakers beter wisten waar ze aan toe waren als zij een nieuw filmplan indienden. Ervaren filmmakers met een grote staat van dienst zouden hun creatieve en innovatieve gang kunnen gaan zonder afhankelijk te zijn van commissies. Daarbij zou een wat matigere film er niet meteen voor moeten zorgen dat deze mensen achteraan in de rij voor subsidie konden gaan staan. Samenwerkingen zouden nieuwe talentvolle filmmakers helpen in de ontwikkeling zonder dat ze daarbij de grote jongens in de weg stonden.

Als dit model werkelijk gebruikt wordt, dan zullen de grote namen makkelijker aan subsidie komen. Zoals aan de resultaten te zien is, zijn de cijfers voor een aantal van hen heel hoog. Dit betekent, voor bijvoorbeeld een producent, dat zeker al de helft van alle punten op de staat van dienst gescoord zou worden. Stel dat één van die mensen met een hoog cijfer een film zou maken die bijna geen bezoekers trekt. In dat geval zal de verwachting van die persoon dalen, maar niet zo veel dat het cijfer keldert. Op deze manier is continuïteit gewaarborgd.

Hoofdstuk 5

Aanbevelingen

Het is belangrijk om genoeg goede data te hebben. De in de discussie genoemde bevindingen zijn gebaseerd op slechts 538 films. Wij raden aan om er voor te zorgen met meer data te werken om de volgende redenen.

Het levert een betere schatter voor σ . Deze is namelijk gebaseerd op alle films die er gemaakt zijn en hoe meer films, hoe beter de schatter is. Ook de verwachtingswaarde voor elke van de filmmakers zal beter geschat kunnen worden. Naast betere schattingen kan er ook gekeken worden naar eventuele andere parameters voor de regressie- en variantieanalyse. We zijn er in de berekening van σ vanuit gegaan dat twee filmmakers onderling onafhankelijk van elkaar zijn. Dit is niet zo als je twee producenten hebt die altijd met elkaar samenwerken. Met een grotere database hadden we mogelijk de covarianties tussen twee filmmakers kunnen berekenen en zo beter hun samenwerking kunnen voorspellen.

De verdisconteringswaarde γ is afhankelijk van de tijd. Deze tijd nemen wij in jaren. Dat betekent dat een film die op 31 december uitgekomen is een heel jaar minder waard is dan een film die een dag later op 1 januari uitkomt. Aangezien veel films in december uitkomen zou het een verschil kunnen maken om de verdiscontering bijvoorbeeld in maanden te nemen.

Hoofdstuk 6

Dankwoord

Op deze plaats willen wij graag iedereen bedanken die op de één of andere manier geholpen heeft met de totstandkoming van dit model en verslag.

Allereerst Judith Vink-Timmer voor de afgelopen 20 weken begeleiden. Niet alleen het wekelijkse verfrissende praatje, maar ook de handvol keren op en neer Amsterdam hebben toch wel wat tijd gekost.

Esmé Lammers, Nathalie Scholten en Marloes Kleiterp voor hun kennis vanuit de filmwereld en het aanleveren van de data.

Richard Boucherie voor de totstandkoming van deze opdracht en zijn inzichten tijdens de koffiepauzes.

Wilbert Kallenberg voor het helpen bedenken van een passend kansmodel en er tot drie keer toe tijd voor maken om het ons uit te leggen.

Tot slot nog een vermelding voor Brend Wanders. Voor het handelbaar maken van de database.

Niek Baër
Maarten Bos
Arjan Feenstra

Bibliografie

- [1] WikiPedia (2008). *De vrije encyclopedie*. <http://www.wikipedia.org>
- [2] IMDbPro (2008). *Internet Movie Database Professional*.
<http://www.imdb.com/>
- [3] Film Fonds 2008. *Nieuwe Selectie Methode*. Vraagstukken, een inleiding. Op aanvraag beschikbaar bij “Het Nederlandse Fonds voor de Film”.
- [4] Hennig-Thurau T., Houston M.B. & Sridhar S. (2006). *Can good marketing carry a bad product? Evidence from the motion picture industry*. *Market Lett*, Vol. 17, 205-219.
- [5] Elberse A. (2007). *The Power of Stars: DO Star Actors Drive the Success of Movies?* *Journal of Marketing*, Vol. 71 (October 2007), 102-120.
- [6] Basuroy S., Chatterjee S. & Ravid S.A. (2003). *How Critical Are Critical Reviews? The Box Office Effects of Film Critics, Star Power, and Budgets*. *Journal of Marketing*, Vol. 67 (October 2003), 103-117.
- [7] King T. (2007). *Does film criticism affect box office earnings? Evidence from movies released in the U.S. in 2003* *J Cult Econ*, Vol. 31, 171-186.

Hoofdstuk 7

Bijlage

Voor het schatten van σ uit alle waarden voor R_{ij} hebben we Maple gebruikt met de onderstaande commando's. Allereerst importeren we de data uit een text-file. Deze datapunten hebben een significantie van twee decimalen. Dit is volgoende significantie om uiteindelijk niet te veel fouten in de cijfers te krijgen.

```
> A:=ImportData{"E:\\UT\\Bachelor\\B0\\Rijv31.txt"};
```

Vervolgens laden we de benodigde bibliotheken in:

```
> with(stats):  
  with(plots):  
  with(stats[statplots]):
```

De R_{ij} die we nodig hebben staan in de 5^e kolom van de matrix. Na deze in een lijst te hebben geplaatst, kan een histogram getekend worden.

```
> Rij:=seq(A[i,5],i=1..538):  
  Rhist:=histogram(Rij,numbars=50):  
  display(Rhist,view=[-10..10,0..0.35]);
```

De steekproefverwachting en steekproefvariantie zijn dan:

```
> describe[mean](Rij);  
  describe[variance](Rij);
```

Vervolgens de verdeling met verwachting 0 en standaardafwijking gelijk aan de steekproefstandaardafwijking:

```
> m:=0:  
  sigmaR:=sqrt(describe[variance](Rij));  
  NR:=exp( -(x-m)^2/2/sigmaR^2 ) / sqrt(2*Pi*sigmaR^2):
```

Het plotten van het histogram samen met de normale verdeling levert het volgende resultaat te zien:

```
> NRplot:=plot(NR,x=-50..50):  
  display({Rhist,NRplot},view=[-10..10, 0..0.35]);
```

In deze Maple commando's maken wij gebruik van data over films verkregen bij het Nederlands Fonds voor de Film. Deze data is vertrouwelijk en is daarom niet opgenomen in dit verslag. Indien men deze data toch wil inzien dient men contact op moeten nemen met het Nederlands Fonds voor de Film.