Inhalt

0. Einleitung

In der vorliegen Arbeit beschäftigen wir uns mit der Bewertung und Ausübung von Optionen. Zu diesem Themenkomplex existiert bereits eine Vielzahl von Publikationen, die sich hauptsächlich mit den Europäischen Optionen beschäftigen und auf Modellen in diskreter Zeit aufbauen. Optionen sind eine Klasse von Wertpapieren, bei denen dem Käufer das Recht eingeräumt wird, ein Wirtschaftsgut (z.B. eine Aktie) zu einem am Kauftag festgelegten Preis innerhalb einer ebenfalls vertraglich festgelegten Frist zu erwerben. Dabei unterscheiden wir zwei verschiedene Arten: die Europäische Option, bei der die Ausübung nur am Tag des Ablaufs der Frist erfolgen kann, und die Amerikanische Option, bei der der Besitzer der Option zu jedem Zeitpunkt innerhalb der Frist die Ausübung vollziehen kann. Da der Preis des zugrundeliegenden Wirtschaftsgutes (Aktie) Schwankungen unterliegt, stellt sich das Problem der Bestimmung des Preises für die Option sowie das der optimalen Ausübung.

Wir werden im ersten Kapitel, basierend auf der Arbeit von IOANNIS KARATZAS [7], ein Modell in stetiger Zeit für die Preisbildung bei Amerikanischen Optionen entwickeln. Dabei gehen wir zur Einstimmung auch den Weg über die Europäischen Optionen und erhalten für diese die bekannte Black und Scholes Formel, die wir in expliziter Form herleiten. Danach wenden wir uns den Amerikanischen Optionen zu, für die wir ebenfalls einen "fairen" Preis bestimmen.

Im zweiten Kapitel wird ein Modell in diskreter Zeit entwickelt, das es uns erlaubt auf Bayessche Weise eine optimale Strategie für die Ausübung Amerikanischen Optionen zu bestimmen. Grundlegend für dieses Kapitel ist die Arbeit von WERNER JAMMERNEGG [6], in der sequentielle binäre Entscheidungen behandelt werden, wobei die Ausübung der Amerikanischen Optionen eine Anwendung darstellt.

An dieser Stelle möchte ich meinem Betreuer Herrn Prof. H.-J. Girlich danken, der mir wichtige und wertvolle Hinweise gab und mir während der Arbeit ständig beratend zur Seite stand.

1. Preisbildung für Amerikanische Optionen

1.1. Das Marktmodell

Wir wollen zuerst einige Annahmen über den Markt, auf dem unsere Wertpapiere bzw. unsere Optionen gehandelt werden, treffen. Wir stellen uns einen Markt vor, auf dem zwei verschiedene Wertpapiere gehandelt werden. Eines dieser ist eine festverzinsliche Anlage, im folgenden mit Bond bezeichnet. Diese hat einen Preis P_0, der folgender Differentialgleichung genügt:

$$\begin{aligned} dP_0(t) &= rP_0(t)dt, \\ P_0(0) &= p_0 = 1 \end{aligned} \qquad (1.1.1)$$

Dieser Preis bestimmt auch unseren Diskontfaktor gemäß

$$\beta(t) := \frac{1}{P_0(t)} = \exp\left\{-\int_0^t r ds\right\} = e^{-rt}, \quad \text{mit } 0 \le t < \infty . \qquad (1.1.2)$$

Das andere Wertpapier ist eine risikobehaftete Anlage, wie zum Beispiel eine Aktie. Deren Preis soll folgender linearen stochastischen Differentialgleichung genügen:

$$\begin{aligned} dP_1(t) &= P_1(t)\big(b(t)dt + \sigma dW_t\big), \quad 0 \le t < \infty \\ P_1(0) &= p_1 > 0 \end{aligned} \qquad (1.1.3)$$

Wobei $W_t, F_t;\ 0 \le t < \infty$ der eindimensionale Standard-Wienerprozeß auf dem Wahrscheinlichkeitsraum $[\Omega, F, P]$ ist und $F_t := \sigma\{W_s,\ 0 \le s \le t\}$ die von W_t erzeugte σ-Algebra bezeichnet.

Der diskontierte Preis βP_1 der risikobehafteten Anlage genügt dann folgender Differentialgleichung:

$$\begin{aligned} d\big[\beta(t)P_1(t)\big] &= P_1(t)d\beta(t) + \beta(t)dP_1(t) \\ &= P_1(t)\big(-r\beta(t)dt\big) + \beta(t)\big(P_1(t)\big[b(t)dt + \sigma dW_t\big]\big) \qquad (1.1.4) \\ &= \beta(t)P_1(t)\big[\big(\beta(t) + r\big)dt + \sigma dW_t\big], \qquad 0 \le t < \infty \end{aligned}$$

In den gesamten folgenden Abschnitten seien

r=const	als Zinsrate
$\{b(t);\ 0 \le t < \infty\}$	als Driftkoeffizient des Preises der risikobehafteten Anlage
σ=const	als Streuungskoeffizient

meßbar, bzw. meßbare Prozesse, adaptiert an $\{F_t\}$ und gleichmäßig beschränkt in $(t, \omega) \in [0, T] \times \Omega$ für alle endlichen T>0. Sie werden als *Koeffizienten des Marktmodells* bezeichnet.

Bemerkung 1.1.1

Der Einfluß des Zufalls erfolgt in unserem Modell also nur über den Wienerprozeß W in 1.1.3 (siehe dazu auch Anhang A1). Die Koeffizienten des Marktmodells sind damit nicht vom Zufall abhängig und werden somit als gegeben und als deterministisch angesehen.

Bemerkung 1.1.2

Die Begriffe meßbarer Prozeß und adaptierter Prozeß werden in Anhang A1 erläutert.

Wir wollen an dieser Stelle noch die Differentialgleichung (1.1.3) lösen. Wir betrachten also folgenden Ausdruck:

$$\frac{dP_1(t)}{P_1(t)} = b(t)dt + \sigma dW_t \,. \qquad (1.1.5)$$

Nach der Integration erhalten wir sofort

$$\int_0^t \frac{dP_1(s)}{P_1(s)} = b(t)t + \sigma W_t \,.$$

Setzen wir nun $g(P_1(t)) := \ln(P_1(t))$ und wenden wir die eindimensionale Itô-Formel (siehe Anhang A2) an, so erhalten wir

$$d(\ln P_1(t)) = \frac{dP_1(t)}{P_1(t)} + \frac{1}{2}\left(-\frac{1}{(P_1(t))^2}\right)(dP_1(t))^2$$

und somit

$$\frac{dP_1(t)}{P_1(t)} = d(\ln P_1(t)) + \frac{1}{2}\sigma^2 dt$$

bzw.:

$$d(\ln P_1(t)) = \frac{dP_1(t)}{P_1(t)} - \frac{1}{2}\sigma^2 dt \,.$$

Nutzen wir nun (1.1.5) und führen die Integration durch, so führt uns dies zu

$$\ln\frac{P_1(t)}{P_1(0)} = \left(b(t) - \frac{1}{2}\sigma^2\right)t + \sigma W_t \,.$$

Daraus folgt nun mit (1.1.3) sofort

$$P_1(t) = p_1\left[\exp\left\{\left(b(t) - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}\right]. \qquad (1.1.6)$$

1.2. Investitions- und Vermögensprozeß

Wir betrachten nun einen "kleinen Investor", d.h. einen Marktteilnehmer, dessen Handlungen auf dem Markt die bestehenden Preise nicht beeinflussen können. Dieser Investor sei mit einem Anfangsvermögen $x \geq 0$ ausgestattet und will dieses nun in die zwei Anlagen unseres Marktes investieren. Wenn er sich entscheidet, $N_i(t)$ Anteile der Anlage i=0,1 zur Zeit t zu besitzen, so berechnet sich sein **Vermögen** X_t zum Zeitpunkt t wie folgt:

$$X_t = \sum_{i=0}^{1} N_i(t)P_i(t), \; 0 \leq t < \infty \qquad (1.2.1)$$

Weiterhin nehmen wir nun an, daß die risikobehaftete Anlage eine **Dividende μ=const** pro Zeiteinheit und pro Geldeinheit abwirft. Wenn die verschiedenen Handelsaktionen auf unserem Markt zu diskreten Zeitpunkten stattfinden, wie zum Beispiel ..., $t-h$, t, $t+h$, ... , so berechnet sich der Zuwachs des Vermögens wie folgt:

$$X_{t+h} - X_t = \sum_{i=0}^{1} N_i(t)\left[P_i(t+h) - P_i(t)\right] + h\mu N_1(t)P_1(t)\,.$$

Gehen wir nun zu dem stetigen Fall über, so folgt sofort

$$dX_t = \sum_{i=0}^{1} N_i(t)dP_i(t) + \mu N_1(t)P_1(t)dt\,. \qquad (1.2.2)$$

Setzen wir nun

$$\pi_i(t) := N_i(t)P_i(t),\ i = 0{,}1,$$

was den Betrag, der in die i-te Anlage investiert wurde, bezeichnet, so erhalten wir für (1.2.2) mit $0 \le t < \infty$ folgende Darstellung:

$$\begin{aligned} dX_t &= \sum_{i=0}^{1} N_i(t)dP_1(t) + \mu\pi_1(t)dt \\ &= N_0(t)rP_0(t)dt + N_1(t)P_1\left[b(t)dt + \sigma dW_t\right] + \mu\pi_1(t)dt \\ &= \left(r\pi_0(t) + b(t)\pi_1(t) + \mu\pi_1(t)\right)dt + \sigma\pi_1(t)dW_t \qquad (1.2.3) \end{aligned}$$

mit $\pi_0(t) = X_t - \pi_1(t)$ gilt dann

$$dX_t = \left[rX_t + \left(b(t) + \mu - r\right)\pi_1(t)\right]dt + \sigma\pi_1(t)dW_t$$

Definition 1.2.1
Ein **Portefeuilleprozeß** $\boldsymbol{\pi} = \left\{\pi_1(t), F_t;\ 0 \le t < \infty\right\}$ ist ein meßbarer, $\mathbb{R}$-wertiger, adaptierter Prozeß, der für alle endlichen T>0 folgender Ungleichung genügt:

$$\int_0^T \pi_1^2(s)ds < \infty \quad P-f.s.$$

$\pi_1(t)$ kann auch negativ werden, was als short-selling oder Leerverkauf bezeichnet wird. Man versteht darunter den Verkauf geliehener Aktien. Auch der Betrag $\pi_0(t) = X_t - \pi_1(t)$ darf negativ werden, was einem Leihen des Geldes zur Zinsrate r entspricht.

Definition 1.2.2
Die eindeutige Lösung der linearen stochastischen Differentialgleichung (1.2.3) ist für alle Portefeuilleprozesse π ist gegeben durch

$$X_t = P_0(t)\left[x + \int_0^t \beta(s)\pi_1(s)\left(b(s) + \mu - r\right)ds + \int_0^t \beta(s)\pi_1(s)\sigma dW_s\right],\ 0 \le t < \infty\,.$$

Sie wird als **Vermögensprozeß bez. des Portefeuilleprozesses $\boldsymbol{\pi}$** bezeichnet.

Wir zeigen jetzt, daß dieses X_t tatsächlich die Lösung von (1.2.3) ist. Bilden wir die Ableitung nach t so folgt daraus:

$$dX_t = P_0(t)\left[\beta(t)\pi_1(t)\left(b(t) + \mu - r\right)dt + \beta(t)\pi_1(t)\sigma dW_t\right] + dP_0\,\frac{1}{P_0(t)}X_t$$

Mit $\beta(t) = \dfrac{1}{P_0(t)}$ und $dP_0(t) = rP_0(t)dt$ erhalten wir sogleich

$$\begin{aligned} dX_t &= \pi_1(t)\left(b(t) + \mu - r\right)dt + \pi_1(t)\sigma dW_t + rX_t dt \\ &= \left[rX_t + \pi_1(t)\left(b(t) + \mu - r\right)\right]dt + \pi_1(t)\sigma dW_t \end{aligned}$$

Definition 1.2.3
Für ein gegebenes Anfangsvermögen x≥0 und einen endlichen Zeithorizont T>0 bezeichnen wir einen Portefeuilleprozeß π als **zulässig auf [0,T] für eine Erstausstattung x≥0** und schreiben $\pi \in A(T,x)$, wenn

$$X_t \geq 0, \quad 0 \leq t \leq T \quad \text{f.s.}$$

gilt. Außerdem führen wir folgende Bezeichnung ein

$$\mathbf{A}(x) := \bigcup_{T>0} \mathbf{A}(T,x)$$

und sagen, daß ein $\pi \in A(x)$ **zulässig für das Anfangsvermögen x≥0** ist.

Diese Definition stellt sicher, daß nur solche Portefeuilleprozesse betrachtet werden, für die das Vermögen des Investors auf ganz [0,T] nichtnegativ bleibt.

Gehen wir zunächst noch einmal kurz auf den Zusammenhang zwischen Aktien- und Bondpreis ein. Aus der Gleichung (1.1.6) sehen wir sofort, daß der verzinsten Aktienpreis $\beta(t)P_1(t) = p_1\left[\exp\left\{\left(b(t) - r - \frac{1}{2}\sigma^2\right)t + \sigma W_t\right\}\right]$ für b(t)<r ein echtes Supermartingal wird und für b(t)>r ein echtes Submartingal, d.h. es gilt

$$E\left[\beta(s)P_1(s)\middle|\mathbf{F}_t\right] < \beta(t)P_1(t)$$

bzw. $$E\left[\beta(s)P_1(s)\middle|\mathbf{F}_t\right] > \beta(t)P_1(t) \text{ für } 0 \leq s < t < \infty.$$

Für die Akteure auf unserem Markt sind jedoch nur Wertpapiere mit $b(t) \geq r$ von Interesse. Nehmen wir an, daß alle Akteure dieselben Einschätzungen für b(t) treffen, so würden im Laufe der Zeit nur die Aktien gekauft, für die gilt $b(t) \geq r$ und nur diejenigen verkauft, für die $b(t) \leq r$ gilt. Dies heißt aber, daß nur Aktien mit b(t)=r wirklich gehandelt würden, d.h. es werden die Aktien gehandelt, für deren diskontierte Preise gilt $\beta(t)P_1(t) = p_1\left[\exp\left\{-\frac{1}{2}\sigma^2 t + \sigma W_t\right\}\right]$. Man bezeichnet einen solchen Markt als risikoneutral. Im folgenden werden wir zeigen, daß sich unser Markt durch eine geschickte Transformation des zugrundeliegenden Wahrscheinlichkeitsmaßes in einen risikoneutralen überführen läßt.

Dazu definieren wir nun den $\mathbb{R}$-wertigen, meßbaren Prozeß

$$\theta_t := \frac{1}{\sigma}\left(b(t) + \mu - r\right), \; F_t, \quad 0 \leq t < \infty, \tag{1.2.4}$$

der in $(t,\omega) \in [0,T] \times \Omega$ für jedes endliche T>0 wegen unserer Forderungen an die Koeffizienten des Marktmodells beschränkt ist. Die Komponenten dieses Prozesses genügen folgender Gleichung:

$$\sigma\theta_t = b(t) + \mu - r, \quad 0 \leq t < \infty, \tag{1.2.5}$$

P-f.s.
Wir definieren weiterhin folgendes exponentielle Supermartingal

$$Z_t := \exp\left\{-\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t |\theta_s|^2 ds\right\}, \; F_{t,} \; 0 \leq t < \infty. \tag{1.2.6}$$

Z ist wirklich ein Supermartingal, denn es gilt für s≤t mit einer Eigenschaft der bedingten Erwartung (siehe Anhang A4) und des Wienerprozesses:

$$E\left[Z_t \middle| \mathbf{F}_s\right] = E\left[\frac{\exp\left\{-\int_0^t \theta_u dW_u\right\}}{\exp\left\{\frac{1}{2}\int_0^t \theta_u^2 du\right\}}\middle| \mathbf{F}_s\right] = E\left[\frac{\exp\left\{-\int_0^s \theta_u dW_u\right\}}{\exp\left\{\frac{1}{2}\int_0^t \theta_u^2 du\right\}}\right] \le E\left[\frac{\exp\left\{-\int_0^s \theta_u dW_u\right\}}{\exp\left\{\frac{1}{2}\int_0^s \theta_u^2 du\right\}}\right] = Z_s$$

Da aufgrund unserer Annahmen für die Koeffizienten unseres Marktmodells gilt, daß $|\theta_t(\omega)|$ lokal beschränkt ist, können wir mit Hilfe der Novikov-Bedingung (siehe dazu Anhang A3) sofort schließen, daß Z_t ein Martingal ist.

Später benötigen wir noch eine andere Darstellung für Z_t, deshalb kommen wir zu folgender

Bemerkung 1.2.4
Z_t besitzt auch die Darstellung

$$Z_t = 1 - \int_0^t Z_s \theta_s dW_s$$

Beweis:

Wir setzen $f(x) := e^x$ und $\eta_t := -\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \theta_s^2 ds$. Wenden wir nun die eindimensionale Itô-Formel (siehe Anhang A2) auf 1.2.6 an, so erhalten wir

$$dZ_t = f'(\eta_t) d\eta_t + \frac{1}{2} f''(\eta_t)\left(d\eta_t\right)^2.$$

Setzen wir nun $d\eta_t = -\theta_t dW_t - \frac{1}{2}\theta_t^2 dt$, $\left(d\eta_t\right)^2 = \theta_t^2 dt$, sowie $f'(x) = e^x = f''(x) = f(x)$ in die vorhergehende Gleichung ein, so erhalten wir

$$dZ_t = -f(\eta_t)\theta_t dW_t$$

und wegen

$$Z_t = Z_0 + \int_0^t dZ_s$$

gilt

$$Z_t = 1 - \int_0^t Z_s \theta_s dW_s \ ,$$

womit wir die gewünschte Darstellung erzielt haben.

•

Mit Hilfe des Girsanov-Theorems (siehe dazu Anhang A3) wissen wir, daß

$$\widetilde{W}_t := W_t - \int_0^t \theta_s ds, \ 0 \le t < \infty \tag{1.2.6}$$

wieder ein eindimensionaler Wienerprozeß ist und können schließen, daß mit (1.2.3) und (1.2.5) folgt:

$$\begin{aligned} dX_t &= \left[rX_t + \pi_1(t)\left(b(t) + \mu - r\right)\right]dt + \pi_1(t)\sigma dW_t \\ &= \left[rX_t + \pi_1(t)\sigma\theta_t\right]dt + \pi_1(t)\sigma dW_t \\ &= rX_t dt + \pi_1(t)\sigma\left[\theta_t dt + dW_t\right] \\ &= rX_t dt + \pi_1(t)\sigma d\widetilde{W}_t \end{aligned} \tag{1.2.7}$$

sowie

$$\beta(t)X_t = x + \int_0^t \beta(s)\pi_1(s)\sigma\left[\theta_s ds + dW_s\right] \\ = x + \int_0^t \beta(s)\pi_1(s)\sigma d\tilde{W}_s \qquad (1.2.8)$$

Aus einer Eigenschaft der bedingten Erwartung folgt:

$$E\left[x + \int_0^s \beta(u)\pi_1(u)\sigma d\tilde{W}_u \middle| F_t\right] = x + \int_0^s \beta(u)\pi_1(u)\sigma d\tilde{W}_u \text{ für } 0 \le s \le t \le T. \qquad (1.2.9)$$

Daher ist die rechte Seite von (1.2.9) ein Martingal, während die linke Seite ein nichtnegativer Prozeß ist, für die gilt

$$\tilde{E}_T\left[\beta(\tau)X_\tau\right] \le \tilde{E}_T\left[X_\tau\right] = x, \ \forall \tau \in S_{0,T}.$$

Bemerkung 1.2.5

Für feste $0 \le u < v \le \infty$ bezeichnet die Menge $S_{u,v}$ die Menge aller Stoppzeiten τ aus $\{F_t\}$ mit Werten in [u,v], und wir schreiben

$$S^* := S_{0,\infty}$$

$$S := \left\{\tau \in S^*;\ \tau < \infty,\ P-\text{f.s.}\right\}$$

Bemerkung 1.2.6

Mit 1.2.6 lauten die Gleichungen für den Aktienpreis bzw. den diskontierten Preis wie folgt:

$$dP_1(t) = P_1(t)\left[b(t)dt + \sigma dW_t\right] \\ = P_1(t)\left[b(t)dt + \sigma\left(d\tilde{W}_t - \theta_t dt\right)\right] \\ = P_1(t)\left[(r-\mu)dt + \sigma d\tilde{W}_t\right] \qquad (1.2.10)$$

bzw. analog:

$$d\left[\beta(t)P_1(t)\right] = \beta(t)P_1(t)\left[(b(t)-r)dt + \sigma dW_t\right] \\ = \beta(t)P_1(t)\left[\sigma d\tilde{W}_t - \mu dt\right] \qquad (1.2.11)$$

Aus (1.2.11) wird offensichtlich, daß der diskontierte Preisprozeß $\{\beta(t)P_1(t), \mathbf{F}_t,\ 0 \le t \le T\}$ für eine Aktie, die keine Dividende abwirft (d.h. μ=0), ein Martingal bez. des Maßes $\tilde{P}_T$ ist.

1.3. Optionen

Wir wollen uns nun den Optionen zuwenden. Dazu schicken wir einige Erläuterungen vorweg. Nehmen wir an, daß wir zum Zeitpunkt t=0 einen Vertrag unterzeichnen, der uns das Recht einräumt, zu einem beliebigen Zeitpunkt τ zwischen t=0 und dem Verfallszeitpunkt t=T eine Aktie zu einem festgelegten (Ausübungs-) Preis C zu kaufen. Wenn nun zum Zeitpunkt t=τ der Preis $P_1(\tau)$ niedriger als der Ausübungspreis ist, so ist der Vertrag für uns wertlos. Wenn nun aber $P_1(\tau) > C$ ist, so können wir unsere Option ausüben, die Aktie zum Preis C erwerben und dann sofort wieder auf dem Markt verkaufen. Wir machen somit einen Nettogewinn von $[P_1(\tau) - C]^+$ Geldeinheiten. Da wir die Vorhersagbarkeit für den Kursverlauf ausschließen, ist τ hier eine Stoppzeit bez. $\{F_t\}$ mit Werten in [0,T].

Ein solcher Vertrag wird als Amerikanische Option bezeichnet, im Gegensatz zu einer Europäischen Option, die die Ausübung nur zum Verfallszeitpunkt, d.h. τ=T, gestattet. Diese beiden Optionsarten sind Finanztitel und werden ihrerseits auf Märkten gehandelt.
Wir wollen uns nun der mathematischen Formulierung des Problems der Bewertung von Optionen widmen.

Definition 1.3.1
Eine **Amerikanische Option (AO) (T,f)** ist ein Finanztitel, bestehend aus:
- (i) einem Verfallszeitpunkt $T \in (0, \infty]$
- (ii) einer Auswahl von Ausübungszeitpunkten $\tau \in S_{0,T}$
- (iii) einer terminalen Auszahlung $f_\tau = [P_1(\tau) - C]^+$

Der Prozeß $F := \{f_t,\ F_t;\ 0 \le t < \infty\}$ sei nichtnegativ, progressiv meßbar und genüge für ein festes μ>0 folgender Bedingung

$$E\left[\sup_{0 \le s \le t} f_s\right]^\mu < \infty$$

für alle $0 < t < \infty$. Weiterhin soll F stetige Pfade besitzen.

Definition 1.3.2
Eine **Europäische Option (EO) (T,F_T)** ist ein Finanztitel, bestehend aus:
- (i) einem Ablaufzeitpunkt $T \in (0, \infty)$ und
- (ii) einer terminalen Auszahlung $f_T = [P_1(T) - C]^+$

Das zentrale Problem ist nun folgendes: **Wie hoch ist ein fairer Preis, der zum Zeitpunkt t=0 für die AO (EO) zu zahlen ist?** Der Einfachheit halber und zur Vorbereitung und Motivation beschäftigen wir uns zunächst mit der Europäischen Option.

Definition 1.3.3
Es seien x≥0 und T>0 zwei gegebene endliche Zahlen. Ein Portefeuilleprozeß $\pi \in A(T, x)$ mit zugehörigem Vermögensprozeß X heißt **Hedging-Prozeß (HP) gegen die Europäische Option**, wenn f.s. gilt:

$$X_T = f_T.$$

Wenn ein solcher Hedging-Prozeß für ein Anfangsvermögen $X_0 = x$ existiert, dann kann der Investor, anstatt die Option zum Zeitpunkt t=0 für den Preis x zu kaufen, auf dem Markt in ein Portefeuille π aus einer festverzinslichen und einer risikobehafteten Anlage investieren, das die Auszahlung f_T der EO dupliziert, d.h. mit diesem Portefeuille kann der Anleger im gleichen Zeitraum genausoviel Gewinn X_T erzielen. Daraus folgt nun, daß der Preis für die EO nicht größer sein darf, als das Anfangsvermögen x.

Definition 1.3.4
Der kleinste Wert für x≥0, für den ein Hedging-Prozeß $\pi \in A(T,x)$ gegen die EO existiert, heißt **fairer Preis** zum Zeitpunkt t=0 für die EO.

Wir betrachten nun den nichtnegativen Prozeß

$$Q_t := \beta(t)f_t,\ F_t,\ \ 0 \le t \le T$$

und bezeichnen mit K_T eine gemeinsame obere Schranke von $\|\theta_t(\omega)\|$ und $\beta(t,\omega),\ (t,\omega) \in [0,T] \times \Omega$. Für jedes endliche α>1 gilt dann offensichtlich

$$Z_t^\alpha = \exp\left\{-\int_0^T \alpha\theta_s dW_s - \frac{1}{2}\int_0^T \|\alpha\theta_s\|^2 ds\right\} \exp\left\{\frac{\alpha(\alpha-1)}{2}\int_0^T \|\theta_s\|^2 ds\right\}$$

und somit

$$EZ_T^\alpha \le \exp\left\{\frac{\alpha(\alpha-1)}{2} TK_T^2\right\} < \infty$$

Mit Hilfe der Hölder-Ungleichung und mit $p := \sqrt{\mu} > 1$ gilt dann

$$\tilde{E}_T\left[\max_{0\le t\le T} Q_t\right]^p \le \left[K_T^{p^2} \cdot E\left[\max_{0\le t\le T} f_t\right]^{p^2}\right]^{\frac{1}{p}} \cdot \left[EZ_T^q\right]^{\frac{1}{q}}. \qquad (1.3.1)$$

Satz 1.3.5
Der faire Preis für eine Europäische Option (T, f_T) wird durch die endliche Zahl

$$\tilde{E}_T(Q_T) = \tilde{E}_T\left[f_T \exp\left\{-\int_0^T r du\right\}\right] \qquad (1.3.2)$$

gegeben.
Weiterhin existiert ein Hedging-Prozeß $\pi \in A(T,x)$, dessen zugehöriger Vermögensprozeß $X = \{X_t,\ F_t,\ 0 \le t \le T\}$ stetig ist und für jedes feste $t \in [0,T]$ der Gleichung

$$X_t = \tilde{E}_T\left[f_T \exp\left\{-\int_t^T r du\right\} \middle| F_t\right] \text{ f.s.} \qquad (1.3.3)$$

genügt.

Beweis:

Für jedes feste x≥0, für das ein Hedging-Prozeß $\pi \in A(T,x)$ existiert, gilt

$$\tilde{E}_T[Q_T] \le x.$$

Deshalb kann der faire Preis nicht kleiner als $\tilde{E}_T[Q_T]$ sein. Es bleibt daher nur noch, einen Hedging-Prozeß mit dieser Zahl als Anfangsvermögen zu finden. Dazu betrachten wir nun den nichtnegativen Prozeß

$$\varsigma_t := P_0(t)\left[\tilde{E}_T[Q_T] + m_t\right],\ 0 \le t \le T, \qquad (1.3.4)$$

wobei m eine rechtsstetige Version des $\tilde{P}_T$-Martingals $\tilde{E}_T[Q_t|\mathbf{F}_t] - \tilde{E}_T[Q_T]$ mit $m_0 = 0$ f.s. ist. Wegen $\tilde{E}_T|Q_T| < \infty$ gilt nach der Bayesschen Formel (siehe dazu Anhang A1):

$$m_t = \frac{1}{Z_t} E\left[Q_T Z_T|\mathbf{F}_t\right] - E\left[Q_T Z_T\right]. \qquad (1.3.5)$$

Wir setzen nun voraus, daß die Pfade von

$$N_t := \tilde{E}_T\left[Q_T Z_T|\mathbf{F}_t\right] \qquad (1.3.6)$$

stetig sind. Daraus folgt nun aus dem Martingaldarstellungssatz (siehe Anhang A1), daß für (1.3.6) folgende Darstellung existiert:

$$N_t = \tilde{E}_T\left[Q_T Z_T\right] + \int_0^t \varphi(s)dW_s,\ 0 \le t \le T, \qquad (1.3.7)$$

f.s., für einen bestimmten, meßbaren und adaptierten Prozeß $\{\varphi(t), \mathbf{F}_t, 0 \le t \le T\}$ mit

$$\int_0^T \varphi^2(s)ds < \infty.$$

Nun hat m jedoch folgende Darstellung

$$m_t = \frac{N_t}{Z_t} - E\left[Q_T Z_T\right]. \qquad (1.3.8)$$

Wir wollen im weiteren nun die Itô-Formel (siehe Anhang A2) anwenden und setzen dazu:

$$m_t = g(t, N_t, Z_t).$$

Es gilt unter Verwendung von $dZ_t = -Z_t \vartheta_t dW_t$ aus Bemerkung 1.2.4:

$$\frac{\partial g}{\partial t} = \frac{N_t}{Z_t}\left(\theta_t dW_t + \frac{1}{2}\theta_t^2 dt\right),$$

$$\frac{\partial g}{\partial N_t} = \frac{\varphi(t)dW_t}{Z_t},$$

$$\frac{\partial g}{\partial Z_t} = \frac{N_t}{Z_t}\left(\theta_t dW_t + \theta_t^2 dt\right),$$

$$\frac{\partial^2 g}{\partial N_t^2} = 0,$$

$$\frac{\partial^2 g}{\partial N_t \partial Z_t} = \frac{\partial^2 g}{\partial Z_t \partial N_t} = \frac{\varphi(t)\theta_t dt}{Z_t}, \text{ sowie}$$

$$\frac{\partial^2 g}{\partial Z_t^2} = \frac{N_t \theta_t^2 dt}{Z_t}.$$

Wenden wir nun die mehrdimensionale Itô-Formel (siehe Anhang A2) an, so erhalten wir

$$\begin{aligned} dm_t &= \frac{1}{Z_t}\left[2N_t\theta_t dW_t + 2N_t\theta_t^2 dt + \varphi(t)dW_t + \varphi(t)\theta_t dt\right] \\ &= \frac{1}{Z_t}\left[\left(2N_t\theta_t + \varphi(t)\right)dW_t + \left(2N_t\theta_t + \varphi(t)\right)dt\right] \end{aligned} \qquad (1.3.9)$$

Wenden wir jetzt die Girsanov-Transformation auf (1.3.9) an, so folgt dann

$$dm_t = \frac{1}{Z_t}\left(2N_t\theta_t + \varphi(t)\right)d\widetilde{W}_t . \qquad (1.3.10)$$

Und wir erhalten für m folgende Darstellung:

$$m_t = \int_0^t \psi(s)d\widetilde{W}_s \;\text{ mit }\; \psi(s) := \frac{1}{Z_t}\left(2N_t\theta_t + \varphi(t)\right). \qquad (1.3.11)$$

Wenn wir nun setzen

$$\pi_1(t) := P_0(t)\frac{\psi(t)}{\sigma},$$

so wird aus (1.3.3) dann (1.2.9) mit der Identifizierung $X \equiv \zeta$ und $x \equiv \widetilde{E}_T\left[Q_T\right]$.

•

Der Satz 1.3.5 ermöglicht es uns, nun den fairen Preis für eine Option zu bestimmen, und wir werden im folgenden Abschnitt auch das bekannteste Ergebnis der Optionsbewertungstheorie mit Hilfe dieses Satzes herleiten.

1.4. Die Black und Scholes Formel

Wir wollen als erstes Resultat die wohlbekannte Black und Scholes Formel für die Bewertung Europäischer Optionen herleiten. Dazu benötigen wir die Lösung folgender Differentialgleichung für den Preis der Aktie:

$$dP_1(t) = P_1(t)\left(\left(r-\mu\right)dt + \sigma d\widetilde{W}_t\right) \qquad (1.4.1)$$

Die Lösung dieser Gleichung lautet wie folgt, wenn wir $f(P_1(t)) = \ln P_1(t)$ setzen und die eindimensionale Itô-Formel darauf anwenden:

$$d\left[\ln P_1(t)\right] = \frac{dP_1(t)}{P_1(t)} + \frac{1}{2}\left(-\frac{1}{P_1^{\,2}(t)}\right)\left(dP_1(t)\right)^2$$

$$\frac{dP_1(t)}{P_1(t)} = d\left[\ln P_1(t)\right] + \frac{1}{2}\sigma^2 dt$$

$$\left(r-\mu\right)t + \sigma\widetilde{W}_1 = \ln\frac{P_1(t)}{P_1(0)} + \frac{1}{2}\sigma^2 t$$

und schließlich erhalten wir

$$P_1(t) = p_1 \cdot \exp\left\{\left(r - \mu - \frac{1}{2}\sigma^2\right)t + \sigma\widetilde{W}_t\right\} \qquad (1.4.2)$$

Nun wenden wir uns der eigentlichen Herleitung der Black und Scholes Formel zu. Nach dem Satz 1.3.5 gilt

$$X_t = \tilde{E}_T\left[f_T \exp\left\{-\int_t^T r du\right\}\middle| F_t\right]$$

$$= \tilde{E}_T\left[\left[P_1(T) - C\right]^+ e^{-r(T-t)}\middle| F_t\right]$$

Für einen Prozeß $P_1(\cdot)$, der zum Zeitpunkt t mit dem Wert $P_1(t)$ startet, gilt:

$$X_t = \tilde{E}_T^{t,P_1(t)}\left[\left[P_1(T) - C\right]^+ e^{-r(T-t)}\right] \qquad (1.4.3)$$

In diese Gleichung setzen wir nun die Lösung (1.4.2) unserer Differentialgleichung (1.4.1) für den Preis der Aktie ein und klammern gleichzeitig $P_1(t)$ aus und erhalten somit:

$$X_t = P_1(t)\cdot \tilde{E}_T^{t,P_1(t)}\left[\left[\exp\left\{\left(r - \mu - \frac{1}{2}\right)(T-t) + \sigma\left(\tilde{W}_T - \tilde{W}_t\right)\right\} - \frac{C}{P_1(t)}\right]^+ e^{-r(T-t)}\right].$$

Wir setzen nun $Z_t := \tilde{W}_T - \tilde{W}_t$ und verwenden, daß Z_t normalverteilt mit $N(0, T-t)$ ist: $X_t = P_1(t)\frac{1}{\sqrt{2\pi}}\int_R \left[\exp\left\{\left(r - \mu - 1/2\,\sigma^2\right)(T-t) + \sigma z\right\} - \frac{C}{P_1(t)}\right]^+ e^{-r(T-t)}\frac{1}{\sqrt{T-t}}\exp\left\{-\frac{1}{2}\frac{z^2}{(T-t)}\right\}dz$ Nimmt der Term in der eckigen Klammer des Integranten einen negativen Wert an, so wird der Integrant identisch Null. Daraus bestimmen wir nun die untere Integrationsgrenze wie folgt:

$$\exp\left\{\left(r - \mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma z\right\} = \frac{C}{P_1(t)}$$

Durch elementares Umformen erhalten wir schnell

$$z = \frac{1}{\sigma}\left(\ln\left(\frac{C}{P_1(t)}\right) - \left(r - \mu - \frac{1}{2}\sigma^2\right)(T-t)\right).$$

Indem wir diese Integrationsgrenze beachten und den obigen Integranten weiter zusammenfassen, nimmt das Integral nun folgende Gestalt an:

$$X_t = P_1(t)\frac{1}{\sqrt{2\pi(T-t)}}\int_{\frac{1}{\sigma}\left(\ln\left(\frac{C}{P_1(t)}\right) - \left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}^{+\infty}\left(\exp\left\{\left(-\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma z\right\} - \frac{C}{P_1(t)}e^{-r(T-t)}\right)\exp\left\{-\frac{1}{2}\frac{z^2}{T-t}\right\}dz$$

Wir wollen diese Integration nun weiter ausführen und betrachten dazu jeden der beiden Summanden im Integranten einzeln:

$$X_t = P_1(t)\frac{1}{\sqrt{2\pi(T-t)}}\int_{\frac{1}{\sigma}\left(\ln\left(\frac{C}{P_1(t)}\right) - \left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}^{+\infty}\exp\left\{\left(-\mu - \frac{1}{2}\sigma^2\right)(T-t) + \sigma z\right\}\exp\left\{-\frac{1}{2}\frac{z^2}{T-t}\right\}dz -$$

$$-\frac{1}{\sqrt{2\pi(T-t)}}\int_{\frac{1}{\sigma}\left(\ln\left(\frac{C}{P_1(t)}\right) - \left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}^{+\infty} C\cdot e^{-r(T-t)}\exp\left\{-\frac{1}{2}\frac{z^2}{T-t}\right\}dz$$

Wir betrachten zuerst den zweiten Summanden und substituieren

$$z := \sqrt{T-t}\cdot x$$

Daraus folgt nun

$$dz = \sqrt{T-t}\cdot dx$$

und für die Integrationsgrenze gilt dann

$$x = \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{C}{P_1(t)}\right) - \left(r - \mu - \frac{1}{2}\sigma^2\right)(T-t)\right)$$

bzw.

$$-x = \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right) + \left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)$$

Verwenden wir nun diese Substitution und Transformation der Integrationsgrenze sowie die Symmetrie der Normalverteilung, so besitzt der zweite Summand folgende Darstellung

$$-e^{-r(T-t)}C\cdot\int_{-\infty}^{\frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right)+\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{1}{2}x^2\right\}dx = e^{-r(T-t)}C\cdot\Phi\left(\frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right)+\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)\right) \tag{1.4.4}$$

Wobei Φ die Verteilungsfunktion der Standard-Normalverteilung ist, gemäß

$$\Phi(u) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{u} e^{-\frac{1}{2}y^2}dy\,.$$

Dieses Resultat merken wir uns und wenden uns dem anderen Summanden zu. Wir werden bei dem ersten Summanden ähnlich vorgehen, wollen uns diesen jedoch noch einmal vor Augen führen:

$$P_1(t)\frac{1}{\sqrt{2\pi(T-t)}}\int_{\frac{1}{\sigma}\left(\ln\left(\frac{C}{P_1(t)}\right)-\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}^{+\infty}\exp\left\{\left(-\mu-\frac{1}{2}\sigma^2\right)(T-t)+\sigma z\right\}\exp\left\{-\frac{1}{2}\frac{z^2}{T-t}\right\}dz =$$

$$= P_1(t)e^{-\mu(T-t)}\frac{1}{\sqrt{2\pi(T-t)}}\int_{\frac{1}{\sigma}\left(\ln\left(\frac{C}{P_1(t)}\right)-\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}^{+\infty}\exp\left\{\left(-\frac{1}{2}\sigma^2\right)(T-t)+\sigma z\right\}\exp\left\{-\frac{1}{2}\frac{z^2}{T-t}\right\}dz$$

Setzen wir wiederum

$$z := \sqrt{T-t}\cdot x$$

Dann erhalten wir analog zum zweiten Summanden:

$$P_1(t)e^{-\mu(T-t)}\frac{1}{\sqrt{2\pi}}\int_{\frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{C}{P_1(t)}\right)-\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}^{+\infty}\exp\left\{\left(-\frac{1}{2}\sigma^2\right)(T-t)+\sigma x\sqrt{T-t}\right\}\exp\left\{-\frac{1}{2}x^2\right\}dx$$

Fassen wir nun die Exponentialausdrücke ihrerseits zusammen, dann ergibt sich folgende Darstellung:

$$\exp\left\{\left(-\frac{1}{2}\sigma^2\right)(T-t)+\sigma x\sqrt{T-t}\right\}\exp\left\{-\frac{1}{2}x^2\right\} = \exp\left\{\left(-\frac{1}{2}\sigma^2\right)(T-t)+\sigma x\sqrt{T-t}-\frac{1}{2}x^2\right\}$$

$$= \exp\left\{-\frac{1}{2}\left(x-\sigma\sqrt{T-t}\right)^2\right\}$$

Setzen wir nun

$$x := u+\sigma\sqrt{T-t}$$

und damit

$$dx=du.$$

Schließlich transformieren wir die Integrationsgrenzen gemäß

$$u = \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{C}{P_1(t)}\right)-\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)-\sigma\sqrt{T-t}$$

$$= \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{C}{P_1(t)}\right)-\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)-\sigma^2(T-t)\right)$$

$$= \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{C}{P_1(t)}\right)-\left(r-\mu+\frac{1}{2}\sigma^2\right)(T-t)\right)$$

bzw.

$$-u = \quad = \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right)+\left(r-\mu+\frac{1}{2}\sigma^2\right)(T-t)\right).$$

Wenden wir nun wiederum die Substitution, die Transformation der Integrationsgrenze und die Symmetrie der Normalverteilung an, so erhalten wir für unseren ersten Summanden:

$$P_1(t)e^{-\mu(T-t)}\frac{1}{\sqrt{2\pi}}\int\limits_{\frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{C}{P_1(t)}\right)-\left(r-\mu-\frac{1}{2}\sigma^2\right)(T-t)\right)}^{+\infty}\exp\left\{\left(-\frac{1}{2}\sigma^2\right)(T-t)+\sigma x\sqrt{T-t}\right\}\exp\left\{-\frac{1}{2}x^2\right\}dx =$$

$$= P_1(t)e^{-\mu(T-t)}\int\limits_{-\infty}^{\frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right)+\left(r-\mu+\frac{1}{2}\sigma^2\right)(T-t)\right)}\frac{1}{\sqrt{2\pi}}\exp\left\{-\frac{u^2}{2}\right\}du$$

$$= P_1(t)e^{-\mu(T-t)}\Phi\left(\frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right)+\left(r-\mu+\frac{1}{2}\sigma^2\right)(T-t)\right)\right) \tag{1.4.5}$$

Fassen wir nun unsere Ergebnisse (1.4.4) und (1.4.5) zusammen, so erhalten wir f.s. $\forall t \in [0,T]$:

$$X_t = \Bigg\{P_1(t)e^{-\mu(T-t)}\Phi\left(\frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right)+\left(r-\mu+\frac{1}{2}\sigma^2\right)(T-t)\right)\right) - Ce^{-r(T-t)}\Phi\Bigg(\frac{1}{\sigma\sqrt{T-t}}\Bigg(\ln\left(\frac{P_1(t)}{C}\right)+\Bigg(r-\mu-\frac{1}{2}\sigma^2$$

und

$$X_T = [P_1(T)-C]^+$$

Wir setzen nun

$$\rho_\pm := \frac{1}{\sigma\sqrt{T-t}}\left(\ln\left(\frac{P_1(t)}{C}\right)+\left(r-\mu\pm\frac{1}{2}\sigma^2\right)(T-t)\right)$$

und erreichen damit das Ziel dieses Abschnitts.

Die **Black und Scholes-Formel**

$$X_t = \begin{cases} P_1(t)e^{-\mu(T-t)}\Phi(\rho_+) - Ce^{-r(T-t)}\Phi(\rho_-) & 0 \le t < T \\ [P_1(T)-C]^+ & t = T \end{cases} \tag{1.4.6}$$

1.5. Die Bewertung Amerikanischer Optionen

Wir haben uns im vorhergehenden Abschnitt mit der Bewertung Europäischer Optionen beschäftigt, mit dem Höhepunkt der Herleitung der Black und Scholes-Formel. In diesem Abschnitt interessieren uns nun die Amerikanischen Optionen, wie sie in Definition 1.3.1 im Abschnitt Optionen definiert wurden. Dazu benötigen wir eine analoge Definition eines Hedging-Prozesses:

Definition 1.5.1
Für einen gegebenen endlichen Zeithorizont T>0 und ein Anfangsvermögen $x \geq 0$ betrachten wir einen Portefeuilleprozeß $\pi \in A(T,x)$ mit seinem zugehörigen Vermögensprozeß X.
π heißt **Hedging-Prozeß** (HP) **gegen die Amerikanische Option (T,f),** und wir schreiben dafür

$$\pi \in H(T,x),$$

wenn für $\tilde{P}_T$-f.a. $\omega \in \Omega$ folgende Bedingungen erfüllt sind:

(i) $X_t(\omega) \geq f_t(\omega), \quad \forall t \in [0,T]$ und

(ii) $X_T(\omega) = f_T(\omega)$

Wir wollen nun noch folgende Bezeichnung einführen:

$$\tau_t := \inf_s \{t \leq s \leq T; X_s = f_s\}$$

$$\tau_t \in S_{t,T}$$

Ein Akteur kauft die Amerikanische Option (T,f) zum Zeitpunkt t=0 zum Preis $x \geq 0$. Nehmen wir weiter an, daß ein HP $\pi \in H(T,x)$ existiert, dann ist es offensichtlich nicht sinnvoll, die AO (T,f) zu einem Zeitpunkt τ $\{0 \leq t \leq T; X_t > f_t\}$ auszuüben, da der Akteur durch eine Investition seines Anfangsvermögens x in das Portefeuille π mehr Gewinn erzielen könnte, als dies durch die Ausübung der AO möglich ist. Ähnlich kann man argumentieren, daß es keinen Sinn macht, die AO zu einem anderen Zeitpunkt als τ_0 in $\{0 \leq t \leq T; X_t = f_t\}$ auszuüben. In diesem Fall sagen wir, daß *der Portefeuilleprozeß π die AO auf [0, τ_0] exakt dupliziert*, d.h. wie bei den Europäischen Optionen können wir ein Portefeuille π bilden, mit dem wir im gleichen Zeitraum den gleichen Gewinn erzielen können. Außerdem erhält das Portefeuille das Vermögensniveau streng über dem der zugehörigen terminalen Auszahlung der AO, d.h.

$$X_t > f_t, \; 0 \leq t < \tau_0 \qquad \text{f.s.}$$

In Analogie zu der Europäischen Option ist der faire Preis für eine AO (T,f) der kleinste Wert für ein Anfangsvermögen $x \geq 0$, das die Erstellung eines HP erlaubt. Wir wollen dies nun genauer definieren:

Definition 1.5.2
Der **faire Preis zur Zeit t=0 für die AO (T,f)** ist die Zahl

$$V_0 := \{x \geq 0; \; \exists \pi \in H(T,x)\} \qquad (1.5.1)$$

Es sei x≥0 irgendeine Zahl, für die ein HP $\pi \in H(T,x)$ existiert. Das Optional Sampling Theorem (siehe dazu Anhang A1) für nichtnegative Supermartingale mit Eigenschaft (i) aus Definition 1.5.1 des HP und

$$\beta(t)X_t = x + \int_0^t \beta(s)\pi_1(s)\sigma d\tilde{W}_s$$

liefert:

$$\tilde{E}_T[Q_\tau] = \tilde{E}_T[\beta(\tau)f_\tau] \le \tilde{E}_T[\beta(\tau)X_\tau] \le x \quad \forall \tau \in S_{0,T}$$

Wir benötigen im weiteren noch folgende Bezeichnungen:

$$u(t) := \sup_{\tau \in S_{t,T}} \tilde{E}_T[Q_\tau] \quad 0 \le t \le T.$$

Und es gilt mit u(0)≤x sowie der Definition 1.5.2 des fairen Preises:

$$u(0) \le V_0. \qquad (1.5.2)$$

Satz 1.5.3
Der faire Preis für eine Amerikanische Option (T,f) zum Zeitpunkt t=0 ist gegeben durch

$$V_0 = u(0) := \sup_{\tau \in S_{0,T}} \tilde{E}_T\left[f_\tau e^{-r\tau}\right]. \qquad (1.5.3)$$

Weiterhin existiert ein Portefeuilleprozeß $\pi \in H(T,x)$ mit zugehörigem Vermögensprozeß $X = \{X_t, F_t:\ 0 \le t \le T\}$, der stetig ist und folgender Gleichung genügt:

$$X_t = \operatorname{ess}\sup_{\tau \in S_{t,T}} \tilde{E}_T\left[f_\tau e^{-r(\tau - t)} \middle| F_t\right] \quad \text{f.s. für jedes feste } t \in [0,T]. \qquad (1.5.4)$$

Bevor wir zu dem Beweis kommen, wollen wir zur Erläuterung noch kurz die Definition von "ess sup" angeben:

$$\operatorname{ess}\sup_{s \in \Omega}(u(s)) := \inf\{K \ge 0 \mid |u(s)| \le K,\ \text{f.ü. in } \Omega\}.$$

Wegen (1.5.2) muß nur die zweite Behauptung gezeigt werden. Bevor wir jedoch den Beweis für diesen Satz vervollständigen können, sind noch einige Vorbereitungen von Nöten.

Wir betrachten nun das Stopp-Problem

$$u(t) := \sup_{\tau \in S_{t,T}} \tilde{E}_T[Q_\tau] \quad 0 \le t \le T.$$

Nach Bismut und Skalli [2], die sich in ihrer Arbeit mit der Theorie des optimalen Stoppens beschäftigen, existiert ein nichtnegatives Supermartingal $Y := \{Y_t, F_t; 0 \le t \le T\}$ mit rechtsseitig stetigen Pfaden und linkem Grenzwert, für den gilt:

$$u(t) = \tilde{E}_T(Y_t) \qquad (1.5.5)$$

$$Y_t = \operatorname{ess}\sup_{\tau \in S_{t,T}} \tilde{E}_T[Q_\tau | F_t] \quad \tilde{P}_T - \text{f.s.} \qquad (1.5.6)$$

$\forall t \in [0,T]$.

Außerdem gilt :

$$Y_0 = u(0)$$

$$Y_T = Q_T \quad \tilde{P}_T\text{-f.s.}$$

Dieser Prozeß ist das kleinste Supermartingal mit rechts-stetigen Pfaden und linksseitigem Grenzwert, das Q majorisiert, und damit ist die Stoppzeit ρ_t gemäß

$$\rho_t := \inf\{t \le s \le T : Y_s = Q_s\}$$

optimal für das obige Stopp-Problem, d.h. es gilt:

$$u(t) = \tilde{E}_T\left[Q_{\rho_t}\right] \quad \forall t \in [0,T].$$

BISMUT und SKALLI [2] zeigen, daß Y regulär ist, d.h. für jede monotone Folge $\{\sigma_n\}_{n=0}^{\infty} \subseteq S_{0,T}$, die $\tilde{P}_T$ -f.s. gegen $\sigma \in S_{0,T}$ konvergiert gilt:

$$\lim_{n\to\infty} \tilde{E}_T\left[Y_{\sigma_n}\right] = \tilde{E}_T\left[Y_\sigma\right].$$

Lemma 1.5.4
Für jedes Supermartingal Y mit (5.5) und (5.6) gilt:
Die Familie $\{Y_\tau\}_{\tau \in S_{0,T}}$ ist gleichmäßig integrierbar bez. $\tilde{P}_T$.

Beweis:

Es gilt $0 \le Y_t \le m_t \ \forall 0 \le t \le T \ \tilde{P}_T - \text{f.s.}$, wobei m eine Modifikation des Martingals $\tilde{E}_T\left[\max_{0\le\vartheta\le T} Q_\vartheta \middle| F_t\right]$ mit rechts-stetigen Pfaden und linkem Grenzwert ist. Mit p>1 und $\frac{1}{p} + \frac{1}{q} = 1$, sowie der Doobschen und der Jensenschen Ungleichung erhalten wir

$$\tilde{E}_T\left[\sup_{0\le t\le T} Y_t\right]^p \le \tilde{E}_T\left[\sup_{0\le t\le T} m_t\right]^p \le q^p \tilde{E}_T\left[m_T{}^p\right] \le q^p \tilde{E}_T\left[\max_{0\le t\le T} Q_t\right]^p,$$

und nach Ungleichung (3.1) ist die letzte Erwartung endlich.

•

Wir haben nun die Voraussetzungen gezeigt, die notwendig sind, um auf Y die DOOB-MEYER-Zerlegung (siehe Anhang A1) anzuwenden. Wir erhalten dann für Y folgende Darstellung:

$$Y_t = u(0) + M_t - \Lambda_t \quad 0 \le t \le T,$$

mit M als $\tilde{P}_T$ -Martingal mit rechts-stetigen Pfaden und linken Grenzwerten;

$$M_0 = \Lambda_0 = 0 \text{ und}$$

$$\tilde{E}_T\left[\Lambda_T\right] = u(0) - \tilde{E}_T\left[Q_T\right].$$

Nach dem Darstellungssatz für Martingale können wir nun für M folgende Darstellung wählen:

$$M_t = \int_0^t \psi(s)d\tilde{W}_s, \quad 0 \le t \le T$$

mit einem meßbaren und adaptierten Prozeß $\{\psi(t), \mathbf{F}_t; 0 \le t \le T\}$ für den gilt:

$$\int_0^T \psi^2(s)ds < \infty \quad \tilde{P}_T - \text{f.s.}$$

Wir können weiter annehmen, daß M stetige Pfade besitzt und somit auch Y.

Wir gelangen nun zum Beweis des Satzes 1.5.3 über den fairen Preis und des zugehörigen HP.

Beweis:

Wir definieren den stetigen und adaptierten Prozeß

$$X_t := \frac{1}{\beta(t)} Y_t, \mathbf{F}_t \quad 0 \le t \le T \qquad (1.5.7)$$

Aus $Y_t = \operatorname{ess\,sup}_{\tau \in S_{t,T}} \tilde{E}_T\left[Q_\tau \middle| F_t\right]$ folgt, daß der Prozeß (5.4) genügt mit

$$X_t(\omega) \ge f_t(\omega) \quad \forall t \in [0,T] \text{ und}$$
$$X_T(\omega) = f_T(\omega)$$

aufgrund der Stetigkeit von F.
Außerdem sind die beiden Stoppzeiten

$$\tau_t = \inf\{t \le s \le T, X_s = f_s\}$$
$$\rho_t = \inf\{t \le s \le T, Y_s = Q_s\}$$

für diese Wahl von X gleich.

Aus $Y_t = u(0) + M_t - \Lambda_t$ und $M_t = \int_0^t \psi(s) d\tilde{W}_s \;\; 0 \le t \le T$ folgt, daß wir X darstellen können gemäß:

$$\beta(t) X_t + \Lambda_t = u(0) + \int_0^t \psi(s) d\tilde{W}_s$$

oder dazu äquivalent in der Form

$$\beta(t) X_t = x + \int_0^t \beta(s)\pi(s)\sigma d\tilde{W}_s \quad 0 \le t \le T$$

mit $\pi(t) := P_0(t)\sigma^{-1}\psi(t)$ und
$\Lambda_t \equiv 0$.

Es zeigt sich, daß X aus (1.5.7) der Vermögensprozeß für das Portefeuille $\pi = \frac{1}{\beta}\frac{1}{\sigma}\psi$ ist, woraus folgt, daß $\pi \in H(T,x)$ gilt.

•

Bemerkung 1.5.5

Lemma 1.5.4 zeigt, daß X gemäß $X_t = \frac{1}{\beta(t)} Y_t, \mathbf{F}_t \;\; 0 \le t \le T$ aus der Klasse D[0,T] (d.h. gleichmäßig integrierbar) bez. $\tilde{P}_T$ ist. Unter dieser Voraussetzung ist es nun möglich zu zeigen, daß der Vermögensprozeß $\vec{X}$ eines Portefeuilles $\vec{\pi} \in \mathbf{H}(T,u(0))$ durch (1.5.4) gegeben und somit eindeutig bestimmt ist.

Wenden wir nämlich das Optional Sampling Theorem (siehe Anhang A1) auf das nichtnegative Supermartingal

$$x + \vec{M}_t := x + \int_0^t \beta(s)\vec{\pi}(s)\sigma d\tilde{W}_s = \beta(t)\vec{X}_t$$

an mit x=u(0), so gilt mit $\vec{X}_t(\omega) \ge f_t(\omega) \;\; \forall t \in [0,T]$:

$$\beta(t)\vec{X}_t \ge \tilde{E}_T\left[\beta(\tau)\vec{X}_\tau \middle| \mathbf{F}_t\right] \ge \tilde{E}_T\left[\beta(\tau) f_\tau \middle| \mathbf{F}_t\right] \quad \tilde{P}_t - \text{f.s.}$$

$\forall \tau \in \mathbf{S}_{t,T}$.

Setzen wir andererseits voraus, daß $\vec{X}$ aus der Familie D[0,T] bez. $\tilde{P}_T$ ist und setzen wir $\vec{\tau}_t := \inf\{t \le s \le T, \vec{X}_s = f_s\}$, dann können wir definieren:

$$\sigma_m := \min\left\{\bar{\tau}_t,\ \inf\left\{t \le s \le T, \int_t^s |\pi(u)|^2 du \ge m\right\}\right\} \quad \forall m \ge 1.$$

Bilden wir nun die bedingte Erwartung bez. $\mathbf{F}_t$ von

$$\beta(\sigma_m)\bar{X}_{\sigma_m} = \beta(t)\bar{X}_t + \int_t^{\sigma_m} \beta(s)\bar{\pi}(s)\sigma d\tilde{W}_s\ ,$$

so erhalten wir

$$\beta(t)\bar{X}_t = \tilde{E}_T\left[\beta(\sigma_m)\bar{X}_{\sigma_m} \middle| \mathbf{F}_t\right]\ \tilde{P}_T - \text{f.s.},$$

wegen $\int_0^T \pi^2(s)ds < \infty$ P-f.s. gilt

$$\lim_{m\to\infty} \sigma_m = \bar{\tau}_t.$$

Mit dem Satz von der monotonen Konvergenz gilt nun aber mit $m \to \infty$:

$$\beta(t)\bar{X}_t = \tilde{E}_T\left[\beta(\bar{\tau}_t)\bar{X}_{\bar{\tau}_t} \middle| \mathbf{F}_t\right]\ \tilde{P}_T \text{-f.s.}$$

Wir können nun schließen, daß $\bar{X}_t$ durch die rechte Seite von (1.5.4) gegeben ist $\forall t \in [0,T]$. Damit ist er nicht von dem Prozeß X aus Satz 1.5.3 zu unterscheiden. Diese Tatsache erlaubt es uns, den Prozeß X als **Bewertungsprozeß für die AO (T,f)** zu bezeichnen.

1.6. Folgerungen aus Satz 1.5.3

Ist Q gemäß $Q_t := \beta(t)f_t,\ \mathbf{F}_t,\ 0 \le t \le T$ ein Submartingal bez. $\tilde{P}_T$, dann folgt mit

$$u(t) = \tilde{E}_T[Y_t] \text{ und}$$

$$Y_t = \operatorname*{ess\,sup}_{\tau \in \mathbf{S}_{t,T}} \tilde{E}_T\left[Q_\tau \middle| \mathbf{F}_t\right]$$

aus dem Optional Sampling Theorem, daß

$$Y_t = \tilde{E}_T\left[Q_T \middle| \mathbf{F}_t\right]\ \tilde{P}_T - \text{f.s.}$$

und $\quad u(t) = \tilde{E}_T[Q_T]$

$\forall t \in [0,T]$ gilt, d.h. $\tau_t = T$ ist optimal.

Es gelte nun $\mu \equiv 0$, dann ist

$$Q_t = \left[\beta(t)P_1(t) - C\beta(t)\right]^+$$

ein Submartingal bez. $\tilde{P}_T$ (vergleiche dazu Bemerkung 1.2.6). Und wir können somit ein Resultat von MERTON [10] angeben:

Folgerung 1.6.1
Eine Amerikanische Option (T,f) mit positivem Ausübungspreis, die auf eine Aktie ohne Dividentenzahlung lautet, sollte nicht vor ihrem Verfallsdatum ausgeübt werden.

D.h., wenn r>0, σ>0 und C>0 ist mit $\mu\equiv 0$, dann ist die Bewertung der AO durch die Black und Scholes-Formel gegeben.

Ist Q gemäß $Q_t := \beta(t) f_t,\ \mathbf{F}_t,\ 0 \le t \le T$ ein Supermartingal bez. $\tilde{P}_T$, dann gilt

$$Y=Q$$

$$u(t) = \tilde{E}_T\left[Q_t\right]$$

und $\tau_t = t$.

Damit liefert die Beziehung

$$X_t = \frac{1}{\beta(t)} Y_t,\ \mathbf{F}_t,\ 0 \le t \le T$$

nun

$$X=f.$$

Es sei nun in unserem Fall C=0 und $\mu\ge 0$, dann ist

$$Q = \beta P_1$$

ein Supermartingal bez. $\tilde{P}_T$ und es gilt $X = P_1$ und wir gelangen nun zur

Folgerung 1.6.2
Besitzt eine Amerikanische Option (T,f) einen Ausübungspreis $C\equiv 0$, so wird sie zu dem gleichen Preis bewertet wie die Aktie, auf die die AO lautet.

Im Falle einer Amerikanischen Option mit $\mu\ge 0$ gilt mit Bemerkung 1.2.6 und

$$X_t = \operatorname*{ess\,sup}_{\tau\in\mathbf{S}_{t,T}} \tilde{E}_T\left[f_\tau e^{-r(\tau-t)} \middle| \mathbf{F}_t\right] \text{ f.s.}$$

nun

$$\begin{aligned}\beta(t)X_t &= \operatorname*{ess\,sup}_{\tau\in\mathbf{S}_{t,T}} \tilde{E}_T\left[\beta(\tau)\left[P_1(\tau) - C\right]^+ \middle| \mathbf{F}_t\right] \\ &\le \operatorname*{ess\,sup}_{\tau\in\mathbf{S}_{t,T}} \tilde{E}_T\left[\beta(\tau)P_1(\tau) \middle| \mathbf{F}_t\right] \\ &\le \beta(t)P_1(t) \quad \tilde{P}_T - \text{f.s.} \quad \forall t \in \left[0,T\right]\end{aligned}$$

und wir gelangen schließlich zur

Folgerung 1.6.3
Die Aktie, auf die die AO lautet, ist mindestens soviel Wert wie die AO selbst, d.h. es gilt

$$X_t \le P_1(t).$$

2. Die optimale Ausübung Amerikanischer Optionen

2.1 Zur Einstimmung

In diesem zweiten Kapitel werden wir uns nun dem Problem der Bestimmung des optimalen Ausübungszeitpunktes für eine Amerikanische Option zuwenden.

Im Gegensatz zum Kapitel 1, in dem wir ein Modell in stetiger Zeit entwickelten, bedienen wir uns hier eines Modells in diskreter Zeit. Aufbauend auf die Arbeit von WERNER JAMMERNEGG [6] wollen wir eine optimale Entscheidungsregel für die Ausübung der Amerikanischen Option herleiten. Wir müssen jedoch in den folgenden Abschnitten erst noch einige theoretische Grundlagen schaffen, die uns den Weg zu unserem Hauptziel ebnen. Dabei ist der Bayessche Satz für Dichten, wie er im Abschnitt 2.4. bewiesen wird, eine der tragenden Säulen unseres stochastischen Entscheidungsmodells.

Gehen wir nun von folgender Situation aus: Der Ertrag bzw. der Gewinn aus einem risikobehafteten Wertpapier (z.B. einer Aktie) wird durch die Wahrscheinlichkeitsverteilung ν_ϑ beschrieben, die von einem unbekanntem Parameter $\vartheta \in \Theta \subseteq R$ abhängt. Nun beobachtet der Entscheidungsträger die Gewinne $(r_1, r_2, ..., r_k)$ im Verlaufe der Zeit. Es erhebt sich jetzt die Frage, ob es nötig ist, alle Werte $(r_1, r_2, ..., r_k)$ für die Bestimmung von ϑ zu benutzen und damit unter Umständen eine gewaltige Menge an Daten zu speichern, oder ob man diese Informationen auf irgendeine Weise derart reduzieren bzw. verdichten kann, ohne daß relevante Informationen bez. des unbekannten Parameters verlorengehen. Diese Überlegung führt uns zu dem Begriff der suffizienten Statistiken bez. eines Parameters für die betreffende Wahrscheinlichkeitsverteilung. Wenn wir vom Bayesschen Standpunkt aus an das Problem herangehen, so hängt die Schätzung des Parameters ϑ nicht nur von den Werten $(r_1, r_2, ..., r_k)$, d.h. also von einer Folge von Realisierungen der reellen Zufallsgröße R_t ab, sondern auch von der A-priori-Verteilung des Parameters ϑ. Diese Verteilung widerspiegelt in gewisser Weise die subjektiven Erwartungen und Eigenschaften des Investors, d.h. der Entscheidungsträger kann seine Erfahrungen und Erwartungen mittels geeigneter Annahmen über diese A-priori Verteilung einfließen lassen. Es wird sich später zeigen, daß mit Hilfe der konjugierten Wahrscheinlichkeitsverteilungen die Berechnung der A-posteriori-Verteilung vereinfacht werden kann, d.h. A-priori- und A-posteriori-Verteilung gehören einer bestimmten Familie von Wahrscheinlichkeitsverteilungen an, nämlich, wie in unserem Fall, der Exponentialfamilie.

2.2. Suffiziente Statistiken

Sei $X = (X_1,...,X_n)$ eine zufällige Stichprobe (SP) vom Umfang n, wobei $X_1,...,X_n$ reelle, unabhängige und identisch verteilte Zufallsvariablen mit der gemeinsamen Verteilung ν_ϑ für festes $\vartheta \in \Theta \subseteq R$ sind. Zur Begriffsbildung sei hier auf Anhang A4 verwiesen. Den Stichprobenraum für die X_i, $1 \le i \le n$ bezeichnen wir mit $\mathfrak{X}$.

Aufgrund des unter Umständen riesigen Speicheraufwandes für die Daten bei wachsendem Stichprobenumfang n, wollen wir eine Statistik T(x) finden, die alle Daten verdichtet und die für uns wichtigen Informationen bez. ϑ enthält. Weiter ist es in unserem Fall wünschenswert, daß die Dimension von T(x) konstant ist, d.h. sie soll sich nicht verändern, auch wenn der Stichprobenumfang sich ändert.

Definition 2.2.1
Sei $X = (X_1,...,X_n)$ eine Stichprobe vom Umfang n mit $X_1,...,X_n$ als unabhängige und identisch verteilte Zufallsvariablen mit gemeinsamer Wahrscheinlichkeitsverteilung ν_ϑ, $\vartheta \in \Theta \subseteq R$. Eine Statistik $Y_n := T(X) = T(X_1,...,X_n)$ heißt **suffizient für ϑ** genau dann, wenn die bedingte Verteilung von X bei gegebenen Y und ϑ nicht vom Parameter $\vartheta \in \Theta$ abhängt.

Beispiel (für Bernoulli-Verteilung):

$P(X_i = 1) = p$ und $P(X_i = 0) = (1-p)$, mit 0<p<1.

Unsere Behauptung ist nun Y_n=T(X)=$\sum_i X_i$ ist suffizient bez. ϑ=p.

Wir müssen zeigen, daß die Verteilung

$$P(X = x \mid \sum X_i = y)$$

unabhängig von p ist.

Es gilt

$$P\left(X = x \mid Y_n = y\right) = \frac{P\left(X = x, Y_n = y\right)}{P\left(Y_n = y\right)}$$

Y_n ist binomialverteilt gemäß

$$P(Y_n = y) = \binom{n}{y} p^y (1-p)^{n-y}$$

Mit $A(T) := \{x : T(x) = T\}$ gilt dann

$$\begin{aligned} P\left(X = x, Y_n = y\right) &= P\left(X_1 = x_1,...,X_n = x_n\right) I_{A(y)}(x) \\ &= \prod_{i=1}^{n} \left(p^{x_i}(1-p)^{1-x_i}\right) I_{A(y)}(x) \\ &= p^{\sum x_i}(1-p)^{n-\sum x_i} I_{A(y)}(x) \\ &= p^y (1-p)^{n-y} \end{aligned}$$

Und damit ist dann

$$P\left(X = x \mid Y_n = y\right) = \frac{1}{\binom{n}{y}}.$$

Dieser Term hängt also nicht mehr von p ab.

Wir bezeichnen im folgenden die Wahrscheinlichkeits-/Dichtefunktion der Wahrscheinlichkeitsverteilung der diskreten/stetigen Zufallsvariablen X mit

$$\nu_\vartheta(x)$$

mit $x = (x_1, ..., x_n) \in X^n$, $\vartheta \in \Theta \subseteq R$.

Da der Nachweis der Suffizienz in den meisten Fällen recht kompliziert ist, wollen wir hier ein einfacheres Kriterium vorstellen. Das Faktorisierungskriterium von Neyman-Fisher gibt eine notwendige und hinreichende Bedingung für die Suffizienz einer Statistik T an:

Satz 2.2.3
Sei $X = (X_1, ..., X_n)$ eine diskrete/stetige Zufallsvariable mit Wahrscheinlichkeits-/Dichtefunktion $\nu_\vartheta(x)$, $x \in X^n$, $\vartheta \in \Theta \subseteq R$.
Eine Statistik T(X) ist suffizient bez. ϑ genau dann, wenn nichtnegative Funktionen g_ϑ, h: $X^n \to R^+$ existieren, so daß $\forall \vartheta \in \Theta$ gilt:

$$\nu_\vartheta(x) = g_\vartheta(T(x))h(x) \qquad (2.2.1)$$

Beweis:

Für den diskreten Fall.
Mit X ist auch T(X) diskret.
$d_\vartheta(x, T(x))$ sei die Verteilungsfunktion von (X,T(X)) und
$g^*_\vartheta(T(x))$ sei die Randverteilung von T(X).

a) T(X) sei suffizient bez. ϑ, für alle T(x)=T, für die gilt $g^*_\vartheta(T(x)) > 0$ gilt dann (gemäß der Definition der Suffizienz), daß die bedingte Verteilung von x unter T(x)=T folgende Gestalt besitzt:

$$\frac{d_\vartheta(x,T)}{g^*_\vartheta(T)} = \begin{cases} w(x) & \text{für } T(x) = T \\ 0 & \text{sonst} \end{cases}$$

wobei $w(x) \geq 0$ von ϑ unabhängig ist.
Dann ist

$$d_\vartheta(x,T) = \begin{cases} w(x)g^*_\vartheta(T) & \text{für } T(x) = T \\ 0 & \text{sonst} \end{cases}$$

Es gilt aber auch

$$d_\vartheta(x,T) = \begin{cases} p_\vartheta(x) & \text{für } T(x) = T \\ 0 & \text{sonst} \end{cases}$$

Daraus folgt dann die Behauptung mit

$$h(x) = w(x)$$

$$g_\vartheta(T(x)) = g^*_\vartheta(T(x))$$

b) Es gelte

$$p_\vartheta(x) = g_\vartheta(T(x))h(x)$$

mit nichtnegativen Funktionen g und h. Die Randverteilung $g^*_\vartheta(T(x))$ von T(X) ist gegeben durch

$$g^*_\vartheta(T(x)) = \sum_{x:T(x)=T} p_\vartheta(x) = p_\vartheta(T(x)) \sum_{x:T(x)=T} h(x)$$

und damit gilt:

$$d_\vartheta(x,T) = \begin{cases} g_\vartheta(T(x))h(x) & \text{für } T(x) = T \\ 0 & \text{sonst} \end{cases}$$

Wir zeigen nun nur noch, daß die bedingte Verteilung von X für T(x)=T unabhängig von ϑ ist.
Für alle T mit $g^* > 0$ gilt:

$$\frac{d_\vartheta(x,T)}{g_\vartheta^*(T)} = \frac{h(x)}{\sum\limits_{x:T(x)=T} h(x)}$$

Der stetige Fall kann in ähnlicher Weise gezeigt werden.

•

Eine Forderung an die Wahrscheinlichkeitsverteilungen, die unsere Marktwerte bzw. Investitionsgewinne kennzeichnet, ist, daß sie zur sogenannten Exponentialfamilie gehören.

Definition 2.2.3
Eine parametrisierte Familie von Wahrscheinlichkeitsverteilungen ist eine (einparametrische) **Exponentialfamilie**, wenn deren Wahrscheinlichkeits- bzw. Dichtefunktion $\bar{v}_\vartheta(x),\ x \in X^n$ für alle $\vartheta \in \Theta$ die folgende Darstellung besitzt:

$$\bar{v}_\vartheta(x) = a(\vartheta)h(x)\exp\{t(x)b(\vartheta)\}$$

wobei a, h, t und b reelle Funktionen sind.

Für die Stichprobe $X = (X_1,...,X_n)$ mit Wahrscheinlichkeits- bzw. Dichtefunktion $f_\vartheta(x_1,...,x_n) = \bar{f}_\vartheta(x_1)\cdot...\cdot\bar{f}_\vartheta(x_n)$ gilt sofort für $x = (x_1,...,x_n) \in X$ und $\vartheta \in \Theta$:

$$f_\vartheta(x) = A(\vartheta)H(x)\exp\{T(x)b(\vartheta)\} \qquad (2.2.2)$$

mit $A(\vartheta) := a(\vartheta)^n,\ H(x) := h(x_1)\cdot...\cdot h(x_n),\ T(x) = t(x_1)+...+t(x_n)$.
Damit sehen wir sofort: Wenn die Wahrscheinlichkeitsverteilung der Zufallsvariablen $X_i,\ 1 \le i \le n$ zur einparametrischen Exponentialfamilie gehört, so gehört auch die Wahrscheinlichkeitsverteilung von $X = (X_1,...,X_n)$ zur einparametrischen Exponentialfamilie.

Folgerung 2.2.4
Wenn die Wahrscheinlichkeitsverteilung der Zufallsvariablen $X_i,\ 1 \le i \le n$ zu einer einparametrischen Exponentialfamilie mit Parameter ϑ gehört, dann ist T(X) in (2.2.2) eine eindimensionale suffiziente Statistik für $\vartheta \in \Theta$.

Wir geben nun einige Beispiele für Verteilungen an, die eine suffiziente Statistik bez. des unbekannten Parameters ϑ besitzen. Wir werden in unserem späteren Modell für die Ausübung der Amerikanischen Optionen vor allem diese Verteilungen betrachten. Sie gehören auch sämtlich der einparametrischen Exponentialfamilie an.

Beispiele 2.2.5:

Wahrschein-lichkeit-/Dichtefunktion	Name bzw. Abkürzung	$\mathfrak{X}$	Θ	(minimale) suff. Statistik	Verteilung von Y_n

$\vartheta^x (1-\vartheta)^{1-x}$	Bernoulli b(1, ϑ)	{0, 1}	(0, 1)	$Y_n = \sum_i X_i$	b(n, ϑ)
$\frac{x^{\alpha-1}\vartheta^{\alpha}}{\Gamma(\alpha)} e^{-\vartheta x}$	Gamma G(α, ϑ)	R^+	R^+	$Y_n = \sum_i X_i$	G(nα, ϑ)
$\frac{1}{\sqrt{2\pi}\sigma} \exp^{-\frac{(x-\vartheta)^2}{2\sigma^2}}$	Normal N(ϑ, σ^2)	R	R	$Y_n = \sum_i X_i$	N(nϑ, nσ^2)
$\frac{1}{\sqrt{2\pi}\sigma x} e^{-\frac{(\log x-\vartheta)}{2\sigma^2}}$	Log-Normal LN(ϑ, σ^2)	R^+	R	$Y_n = \prod_i X_i$	LN(nϑ, nσ^2)

2.3. Der Bayessche Satz für Dichten

Für unsere weiteren Betrachtungen benötigen wir nun bedingte Wahrscheinlichkeiten und bedingte Dichten, die wir nun einführen wollen. Doch zuerst möchten wir die allgemein bekannte Definition für die bedingte Wahrscheinlichkeit angeben:

Definition 2.3.1
In einem Wahrscheinlichkeitsraum $[\Omega, F, P]$ seien zwei Ereignisse $A, B \in F$ gegeben, mit P(B)>0. Die **bedingte Wahrscheinlichkeit** des Ereignisses A unter der Bedingung B wird mit $P(A|B)$ bezeichnet und es gilt

$$P(A|B) = \frac{P(A \cap B)}{P(B)}.$$

Es gilt nun folgender

Satz 2.3.2
Auf einem Wahrscheinlichkeitsraum $[\Omega, F, P]$ seien beliebige Ereignisse $A_k \in F$ mit $P\left\{\bigcap_{k=1}^{l-1} A_k\right\} \neq 0$ gegeben. Es gilt dann

$$P\left\{\bigcap_{k=1}^{l} A_k\right\} = P\{A_1\} P(A_2|A_1) \cdot \ldots \cdot P\left(A_l \middle| \bigcap_{k=1}^{l-1} A_k\right).$$

Beweis:

Wir ersetzen die bedingten Wahrscheinlichkeiten auf der rechten Seite durch ihre Definition als Quotienten und erhalten durch anschließendes Kürzen:

$$P\{A_1\} P(A_2|A_1) \cdot \ldots \cdot P\left(A_l \middle| \bigcap_{k=1}^{l-1} A_k\right) = P\{A_1\} \frac{P(A_1 \cap A_2)}{P(A_1)} \cdot \frac{P\left(\bigcap_{k=1}^{3} A_k\right)}{P(A_1 \cap A_2)} \cdot \ldots \cdot \frac{P\left(\bigcap_{k=1}^{l-1} A_k\right)}{P\left(\bigcap_{k=1}^{l-2} A_k\right)} \cdot \frac{P\left(\bigcap_{k=1}^{l} A_k\right)}{P\left(\bigcap_{k=1}^{l-1} A_k\right)}$$

$$= P\left(\bigcap_{k=1}^{l} A_k\right)$$

•

Aus der Definition der bedingten Wahrscheinlichkeit folgt nun sehr leicht der Satz von der totalen Wahrscheinlichkeit:

Satz 2.3.3
Es sei ein Wahrscheinlichkeitsraum $[\Omega, F, P]$ mit Ereignissen $A, C_k \in F \quad k \in N$ gegeben. Es gelte weiterhin $\bigcup_k C_k = \Omega$ und für $i \neq j$: $C_i \cap C_j = \varnothing$. Dann gilt

$$P\{A\} = \sum_i P(A|C_i) P(C_i)$$

Beweis:

Weil die Ereignisse C_k ein vollständiges Ereignissystem bilden, gilt:

$$A = A \bigcup_k C_k = \bigcup_k \left(A \cap C_k\right)$$

Da nun die C_k paarweise disjunkte Ereignisse sind, gilt

$$P(A) = \sum_k P\left(A \cap C_k\right) = \sum_k P\left(A \cap C_k\right)\frac{P(C_k)}{P(C_k)} = \sum_k P\left(A \middle| C_k\right)P(C_k)$$

•

Aus der Definition 2.3.1 und dem Satz 2.3.3 von der totalen Wahrscheinlichkeit folgt nun die bekannte

Bayessche Formel
Unter den Voraussetzungen des letzten Satzes gilt

$$P\left(C_j \middle| A\right) = \frac{P\left(A \middle| C_j\right)P\left(C_j\right)}{\sum_k P\left(A \middle| C_k\right)P\left(C_k\right)}, \quad \forall A, C_j \in F \tag{2.3.1}$$

Beweis:

Unter Verwendung der Definition der bedingten Wahrscheinlichkeit und des Satzes von der totalen Wahrscheinlichkeit gilt folgendes:

$$P\left(C_j \middle| A\right) = \frac{P\left(A \cap C_j\right)}{P(A)} = \frac{P\left(A \middle| C_j\right)P\left(C_j\right)}{\sum_k P\left(A \middle| C_k\right)P\left(C_k\right)}$$

•

Wir wollen im folgenden die Bayessche Formel für Dichten von Verteilungen herleiten. Dazu führen wir die gemeinsame Verteilung zweier Zufallsgrößen X und Θ auf unserem Wahrscheinlichkeitsraum $[\Omega, F, P]$ ein. Der Zufallsvektor Z:=(X,Θ) besitzt dann die Verteilungsfunktion

$$F_Z(x,\vartheta) := P\left\{X \le x, \Theta \le \vartheta\right\}$$

mit der zugehörigen Dichte $f_Z(x,\vartheta)$, d.h. es gilt

$$F_Z(x,\vartheta) = \int_{-\infty}^{x}\int_{-\infty}^{\vartheta} f_Z(u,v)\cdot du \cdot dv \quad \forall x, \vartheta \in R \ .$$

Die Dichten der Verteilungen für X bzw. für Θ ergeben sich dann als folgende Randverteilungen:

$$f_X(\cdot) := \int_{-\infty}^{+\infty} f_Z(\cdot,\vartheta)d\vartheta$$

bzw.

$$f_\Theta(\cdot) := \int_{-\infty}^{+\infty} f_Z(x,\cdot)dx \, .$$

Wir benötigen nun den Begriff der bedingten Dichte und bringen deshalb die folgende Definition.

Definition 2.3.4
Es sei Z=(X,Θ) ein stetiger Zufallsvektor. Die Funktion $f\left(x \middle| \vartheta\right)$ heißt dann **bedingte Dichte** der Zufallsgröße X unter der Bedingung Θ=ϑ, falls $\forall x, \vartheta \in R$ gilt:

$$f_Z(x,\vartheta) = f\left(x \middle| \vartheta\right) f_\Theta(\vartheta)$$

Wir sind nun in der Lage, die Bayessche Formel für Dichten anzugeben:

Bayessche Formel für Dichten
Für $f_X(\cdot) > 0$ gilt

$$f(\vartheta|x) = \frac{f(x|\vartheta)f_\Theta(\vartheta)}{\int_{-\infty}^{+\infty} f(x|\vartheta)f_\Theta(\vartheta)d\vartheta}. \qquad (2.3.2)$$

Beweis:

Nach der Definition der bedingten Dichte durch Einsetzen der entsprechenden Randverteilungen gilt sofort:

$$f(\vartheta|x) = \frac{f_Z(x,\vartheta)}{f_X(x)} = \frac{f(x|\vartheta)f_\Theta(\vartheta)}{\int_{-\infty}^{+\infty} f(x|\vartheta)f_\Theta(\vartheta)d\vartheta}.$$

•

Wir bezeichnen die Randverteilungsdichten auf der rechten Seite im folgenden als **A-priori-Dichte** und die bedingte Dichte auf der linken Seite der Gleichung als **A-posteriori-Dichte**.

Für $E[|\Theta|] < \infty$ bezeichnet man

$$\int_{-\infty}^{+\infty} \vartheta f(\vartheta|x)d\vartheta =: E[\Theta|x]$$

auch als **bedingte Erwartung** der Zufallsgröße Θ für {X=x}. Einige Eigenschaften der bedingten Erwartung findet man in Anhang A4.

Beispiel 2.3.5:
Es sei nun X normalverteilt, d.h. $N(\vartheta,\sigma^2)$, mit unbekanntem Parameter $\vartheta \in \Theta \subseteq R$. Die A-priori-Verteilung π von J sei auch eine Normalverteilung, nämlich $N(\mu,\tau^2)$ mit bekannten Parametern μ und τ^2. Dann gilt für die gemeinsame Verteilungsdichte h(x,ϑ):

$$h(x,\vartheta) = f(x|\vartheta)\pi(\vartheta) = \frac{1}{2\pi\sigma\tau}\exp\left\{-\frac{1}{2}\left(\frac{[\vartheta-\mu]^2}{\tau^2} + \frac{[x-\vartheta]^2}{\sigma^2}\right)\right\}$$

Um die Randverteilungsdichte $f_X(x)$ zu bestimmen, definieren wir nun:

$$\rho := \frac{1}{\tau^2} + \frac{1}{\sigma^2} = \frac{\sigma^2+\tau^2}{\sigma^2\tau^2}$$

Die quadratische Ergänzung liefert dann folgendes:

$$\frac{1}{2}\left(\frac{[\vartheta-\mu]^2}{\tau^2} + \frac{[x-\vartheta]^2}{\sigma^2}\right) = \frac{1}{2}\left[\left(\frac{1}{\tau^2}+\frac{1}{\sigma^2}\right)\vartheta^2 - 2\left(\frac{\mu}{\tau^2}+\frac{x}{\sigma^2}\right)\vartheta + \left(\frac{\mu^2}{\tau^2}+\frac{x^2}{\sigma^2}\right)\right]$$

$$= \frac{1}{2}\rho\left[\vartheta^2 - \frac{2}{\rho}\left(\frac{\mu}{\tau^2}+\frac{x}{\sigma^2}\right)\vartheta\right] + \frac{1}{2}\left(\frac{\mu^2}{\tau^2}+\frac{x^2}{\sigma^2}\right)$$

$$= \frac{1}{2}\rho\left[\vartheta - \frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\vartheta\right]^2 - \frac{1}{2\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)^2 + \frac{1}{2}\left(\frac{\mu^2}{\tau^2} + \frac{x^2}{\sigma^2}\right)$$

$$= \frac{1}{2}\rho\left[\vartheta - \frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\right]^2 + \frac{(\mu - x)^2}{2(\sigma^2 + \tau^2)}$$

Daraus folgt nun:

$$h(x,\vartheta) = \frac{1}{2\pi\sigma\tau}\exp\left\{-\frac{1}{2}\rho\left[\vartheta - \frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\right]^2\right\}\exp\left\{-\frac{(\mu - x)^2}{2(\sigma^2 + \tau^2)}\right\}$$

und

$$m(x) = \int_{-\infty}^{+\infty} h(x,\vartheta')d\vartheta' = \frac{1}{\sqrt{2\pi\rho}}\frac{1}{\sigma\tau}\exp\left\{-\frac{(\mu - x)^2}{2(\sigma^2 + \tau^2)}\right\}$$

Schließlich gilt dann:

$$\pi(\vartheta|x) = \frac{h(x,\vartheta)}{m(x)} = \sqrt{\frac{\rho}{2\pi}}\exp\left\{-\frac{1}{2}\rho\left[\vartheta - \frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right)\right]^2\right\}$$

Beachte jedoch, daß die Randverteilung von X eine $N(\mu, \sigma^2 + \tau^2)$-Verteilung ist und die A-posteriori-Verteilung des Parameters J bei gegebenen x eine $N\left(\mu(x), \frac{1}{\rho}\right)$-Verteilung ist mit $\mu(x) = \frac{1}{\rho}\left(\frac{\mu}{\tau^2} + \frac{x}{\sigma^2}\right) = \frac{\sigma^2}{\sigma^2 + \tau^2}\mu + \frac{\tau^2}{\sigma^2 + \tau^2}x = x - \frac{\sigma^2}{\sigma^2 + \tau^2}(x - \mu)$.

Wir können auch weiterhin zeigen, daß diese Normalverteilung bei ständiger Beobachtung gegen eine Einpunktverteilung strebt:

Dabei interessieren wir uns hauptsächlich für die Varianz, denn wenn wir das Verschwinden der Varianz zeigen, so strebt diese Verteilung gegen eine Einpunktverteilung.

Es sei nun ϑ A-priori verteilt nach $N(\mu, \tau_0^2)$ und die Verteilung für x sei wiederum $N(\vartheta, \sigma^2)$. Die obigen Überlegungen haben gezeigt, daß die Varianz τ_1^2 nach der Realisierung $X = x_1$ die Gestalt hat:

$$\tau_1^2 = \frac{\sigma^2\tau_0^2}{\sigma^2 + \tau_0^2}.$$

Iterieren wir obiges Herangehen, so folgt für die nächsten Realisierungen $X = x_2$ und $X = x_n$:

$$\tau_2^2 = \frac{\sigma^2\tau_1^2}{\sigma^2 + \tau_1^2} = \frac{\sigma^2\frac{\sigma^2\tau_0^2}{\sigma^2 + \tau_0^2}}{\sigma^2 + \frac{\sigma^2\tau_0^2}{\sigma^2 + \tau_0^2}}$$

$$= \frac{\frac{\sigma^2\tau_0^2}{\sigma^2 + \tau_0^2}}{1 + \frac{\tau_0^2}{\sigma^2 + \tau_0^2}} = \frac{\sigma^2\tau_0^2}{\sigma^2 + 2\tau_0^2}$$

$$\tau_3^2 = \frac{\sigma^2\tau_2^2}{\sigma^2+\tau_2^2} = \frac{\sigma^2\tau_0^2}{\sigma^2+3\tau_0^2}$$ und schließlich:

$$\tau_n^2 = \frac{\sigma^2\tau_{n-1}^2}{\sigma^2+\tau_{n-1}^2} = \frac{\sigma^2\tau_0^2}{\sigma^2+n\tau_0^2}\,.$$

Wir sehen daraus sofort, daß für $n \to \infty$ dieser Ausdruck verschwindet, was wir zeigen wollten. Daraus folgt, daß bei fortdauernder Beobachtung der Realisierungen der Zufallsgröße X die A-posteriori-Verteilung des unbekannten Parameters gegen eine Einpunktverteilung mit der Masse im "wahren" Parameter, der der Verteilung zugrunde liegt, strebt.
Die A-posteriori-Verteilung des Parameters ist nach n Beobachtungen dann die Normalverteilung $N\left(\mu(n,y),\tau_n^2\right)$. Wir wollen nun noch den Parameter $\mu(n,y)$ nach n Beobachtungen berechnen, wobei $y_n = x_1+...+x_n$ ist.
Nach der ersten Beobachtung galt

$$\mu(1,y_1) = \mu(y_1) = \frac{1}{\rho}\left(\frac{\mu}{\tau_0^2}+\frac{y_1}{\sigma^2}\right) = \frac{\sigma^2\mu+\tau_0^2 y_1}{\sigma^2+\tau_0^2}$$

Iterieren wir dieses Verfahren, so erhalten wir

$$\mu(2,y_2) = \frac{\sigma^2\mu(1,y_1)+\tau_1^2 x_2}{\sigma^2+\tau_1^2} = \frac{\sigma^2\,\dfrac{\sigma^2\mu+\tau_0^2 y_1}{\sigma^2+\tau_0^2} + \dfrac{\sigma^2\tau_0^2}{\sigma^2+\tau_0^2}\,x_2}{\sigma^2+\dfrac{\sigma^2\tau_0^2}{\sigma^2+\tau_0^2}}$$

$$= \frac{\sigma^2\mu+y_1\tau_0^2+x_2\tau_0^2}{\left(\sigma^2+\tau_0^2\right)\left(1+\dfrac{\tau_0^2}{\sigma^2+\tau_0^2}\right)}$$ und mit $y_n = x_1+...+x_n = y_{n-1}+x_n \quad \forall n \geq 1$

gilt $$= \frac{\sigma^2+y_1\tau_0^2+x_2\tau_0^2}{\sigma^2+\tau_0^2+\tau_0^2} = \frac{\sigma^2+y_2\tau_0^2}{\sigma^2+2\tau_0^2}$$

und allgemein für n Beobachtungen:

$$\mu(n,y_n) = \frac{\sigma^2\mu+\tau_0^2 y_n}{\sigma^2+n\tau_0^2}\,.$$

2.4. Konjugierte A-posteriori-Verteilungen

Vom rechnerischen Standpunkt aus ist es natürlich wünschenswert, daß sowohl die A-posteriori-Verteilung μ für den Parameter, als auch die A-posteriori-Verteilung p für die Zufallsvariable X Mitglieder einer bekannten Familie von Wahrscheinlichkeitsverteilungen sind. Dies sollte für jede Stichprobeninformation, ausgedrückt durch (n,y), gelten.

Definition 2.4.1
Wenn die A-priori-Verteilung $\bar{\mu}$ und die A-posteriori-Verteilung μ des Parameters ϑ für alle (n,y) zu derselben Familie von Verteilungen gehören, so nennen wir dies eine **(natürliche) konjugierte Familie von Verteilungen**.

Diese Bezeichnung drückt die spezielle funktionelle Beziehung zwischen den Verteilungsfamilien der Zufallsvariablen X und der des Parameters aus.

Ein konkretes Beispiel ist auch schon bei der obigen Herleitung der A-posteriori-Verteilung bei normalverteiltem X und für nomalverteiltes ϑ angegeben worden. Eine Zusammenfassung finden wir in folgender Tabelle. In der vierten Spalte dieser Tabelle sehen wir auch den Typ der A-posteriori-Verteilung p der Zufallsvariablen X.

Beispiele 2.4.2

W-Verteilung ν_ϑ	A-priori-Verteilung $\bar{\mu}$	a-posteriori Vert. μ	a-posteriori Vert. p
Bernoulli $b(1,\vartheta)$	Beta $B(\gamma,\delta)$	Beta $B(\gamma+y,\delta+n-y)$	b(1,p(n,y)) $p(n,y):=\frac{\gamma+y}{\gamma+\delta+n}$
Gamma $G(\alpha,\vartheta)$	Gamma $G(\gamma,\delta)$	Gamma $G(\gamma+n\alpha,\delta+y)$	Pearson-Typ VI
Normal $N(\vartheta,\sigma^2)$	Normal $N(\mu,\tau^2)$	Normal $N(\mu(n,y),\tau_n^2)$ $\mu(n,y):=\frac{\tau^2 y+\sigma^2\mu}{\tau^2 n+\sigma^2}$	Normal $N(\mu(n,y),\sigma^2+\tau_n^2)$ $\tau_n^2:=\frac{\sigma^2\tau^2}{\tau^2 n+\sigma^2}$
Lognormal $LN(\vartheta,\sigma^2)$	Normal $N(\mu,\tau^2)$	Normal $N(\mu_l(n,y),\tau_n^2)$ $\mu_l(n,y):=\frac{\tau^2\log\{y\}+\sigma^2}{\tau^2 n+\sigma^2}$	Lognormal $LN(\mu_l(n,y),\sigma^2+\tau_n^2)$

Bei Wahl derartiger Verteilungen für unsere Marktwerte bzw. für die unbekannten Parameter wird es also relativ leicht sein, die A-posteriori-Verteilungen zu bestimmen.

2.5. Das Modell zur Ausübung Amerikanischer Optionen

Wir kommen nun zu dem Hauptgegenstand dieses Kapitels, nämlich zum Herleiten eines Modells, das uns gestattet, den optimalen Zeitpunkt für die Ausübung einer Amerikanischen Option zu bestimmen. Doch zuvor wollen wir noch einen Vektor einführen, der es uns gestattet, die Informationen, die uns bis zum Zeitpunkt t=n zur Verfügung stehen zu beschreiben.

Den Vektor $(k,b) \in Z \times R^{+}$ bezeichnen wir als Zustandsinformation zur Zeit t=k-L, wobei der Zeitpunkt t=0 den Zeitpunkt des Erwerbs unserer Option darstellt. Die Größe b ist der Wert unserer suffizienten Statistik bezüglich der Marktwerte zum Zeitpunkt t=k-L. Wie zu sehen ist, beginnen wir jedoch schon vor dem Zeitpunkt t=0, nämlich zur Zeit $t=-L,\ L \geq 0$, mit der Sammlung von Informationen über den Kursverlauf unserer zugrundeliegenden Aktie. Der Zeitpunkt t=T markiert den Verfallszeitpunkt, d.h. den Zeitpunkt, an dem unsere Option abläuft. Wir haben bei einer Amerikanischen Option also das Recht erworben, innerhalb der Zeitspanne beginnend mit t=0 und endend mit t=T eine Aktie zu einem am Kaufzeitpunkt festgelegten Preis C zu kaufen, d.h. zu jedem Zeitpunkt $t \in \{1,2,3,...,n,...,T\}$.

Wir wenden uns nun dem Marktwert unserer Option zu. Dazu nehmen wir an, daß dieser beschrieben wird durch die reelle Zufallsgröße $R_t \in R$, $t \in \{-L,-L+1,...,0,...,T-1,T\}$. Diese R_t nehmen also zu jedem Zeitpunkt t Werte aus dem selben Zustandsraum $X = R$ an. Wir nehmen im weiteren an, daß diese Zufallsgrößen R_t verteilt sind nach einer Wahrscheinlichkeitsverteilung mit der Dichte ν_ϑ, die von einem unbekanntem Parameter $\vartheta \in \Theta \subseteq R$ abhängt. Wir beschränken uns im weiteren bei der Wahl von Verteilungen auf die stetigen aus dem Beispiel 2.4.2 und nehmen an, daß die Marktwerte eine unabhängige und identisch verteilte Folge von Zufallsvariablen bilden. Für diese existieren dann suffiziente Statistiken $b = T(r_{-L},...,r_t) = Y_t$ jeweils zu Beginn einer Periode $t,\ 1 \leq t \leq T+1$. Daraus folgt nun, daß die oben eingeführte Zustandsinformation zu Beginn der Periode t gegeben ist durch (k,b) mit k=L+1+t. Aus Beispiel 2.2.6 wird ersichtlich, daß nach dem Übergang von k nach k+1 die Zustandsinformation wie folgt aussieht:

$$(k,b) \to (k+1, v(b,r)) \text{ mit } v(b,r) = \begin{cases} b+r & \text{für Bernoulli-, Gamma-} \\ & \text{und Normalverteilung} \\ b \cdot r & \text{für Log-Normalverteilung} \end{cases}$$

Wenn wir nun den unbekannten Parameter $\vartheta \in \Theta \subseteq R$ betrachten und annehmen, er sei verteilt nach der A-priori-Verteilung mit der Dichte $\overline{\pi}$, dann läßt sich mit Hilfe des Bayesschen Satzes für Dichten die A-posteriori-Verteilung $\pi = \pi(k,b;\cdot)$ für den Parameter $\vartheta \in \Theta \subseteq R$ angeben, nämlich

$$\pi(k,b;D) = \frac{\int_D g(k,b;\vartheta)\overline{\pi}(d\vartheta)}{\int_\Theta g(k,b;\vartheta')\overline{\pi}(d\vartheta')}, \quad D \subseteq \Theta,$$

wobei $g(k,b;\cdot)$ die Dichtefunktion unserer suffizienten Statistik b ist. Mit Hilfe dieser A-posteriori-Verteilung und der Anwendung des Bayesschen Satzes können wir nun die A-posteriori-Übergangswahrscheinlichkeitsdichte $q_t = q_t(k,b;\cdot)$ für unsere Marktwerte zum Zeitpunkt t angeben:

$$q_t(k,b;G) = \int_\Theta \nu_\vartheta(G)\pi(k,b;d\vartheta)\ ,\ G \subseteq X\ .$$

Weiterhin haben wir zu jedem Zeitpunkt $t \in \{1,2,...,T\}$ die Möglichkeit, zu entscheiden, ob wir die Option ausüben oder nicht. Das bedeutet, wir haben zwei alternative Aktionen zur Verfügung, und deshalb führen wir nun die Aktionenmenge $\{A(x),\ x \in X\}$ ein mit

$$A(x) = A = \{0,1\},\ \forall x \in X$$

$$a = \begin{cases} 0: & \text{die Option wird nicht ausgeübt} \\ 1: & \text{die Option wird ausgeübt} \end{cases}$$

Die Gewinn- bzw. Nutzenfunktion für den Zeitpunkt $1 \le t \le T$ sei J_t. Sie stellt den gesamten erwarteten Nutzen ab dem Zeitpunkt t bis zum Verfall der Option dar. Wenn wir nun noch die Endnutzenfunktion $J_{T+1} \equiv 0$ einführen, da bei t=T+1 unsere Option abgelaufen ist und somit keinen Nutzen mehr erbringen kann, so erhalten wir unser

stochastisches Entscheidungsmodell
Für die optimale Ausübung unserer Amerikanischen Option nutzen wir folgendes stochastisches Entscheidungsmodell:

$$M = \left[\left(X, \{0,1\}, q_t(k,b;\cdot), \mathbf{J}_t\right)_{t=1,...,T}, X, \mathbf{J}_{T+1}\right],$$

wobei wir die Bezeichnung der einzelnen Komponenten von oben übernehmen.

Für die weiteren Betrachtungen benötigen wir noch den Begriff des A-posteriori-Barwertes N(k,b) unserer Option, dieser berechnet sich zum Beginn einer jeden Periode t, $1 \le t \le T$, wie folgt:

$$N(k,b) := \int rq(k,b;dr) - C\ .$$

Wie zu sehen ist, stellt er den erwarteten Marktwert bezüglich der A-posteriori-Übergangswahrscheinlichkeit, vermindert um den Ausübungspreis C dar. Er wird bei der Suche nach einem optimalen Ausübungszeitpunkt eine wichtige Rolle spielen.

Da die Nutzenfunktionen J_t den gesamten Nutzen von t an bis zum Verfall darstellen, uns jedoch der erwartete Gesamtnutzen interessiert, benötigen wir für die Formulierung unseres Entscheidungsproblems noch den Begriff des maximalen bedingt erwarteten Gesamtnutzens.

Definition 2.5.1
Der **maximale bedingt erwartete Gesamtnutzen** von der Periode k an bis zum Verfall der Option, unter der Bedingung, daß die Vorgeschichte (k,b) als Zustandsinformation gegeben ist, wird definiert durch

$$\mathbf{J}_t^*(k,b) := \sup_{a \in A}\{\mathbf{J}_t(k,b,a)\}\ .$$

Nun sind wir in der Lage, das Entscheidungsproblem zu formulieren:

Gesucht ist eine Folge

$$f = \left(f_1, f_2, \ldots, f_T\right)$$

von Entscheidungsfunktionen mit

$$f_i(k,b) \in A \ ,$$

die unseren Gesamtnutzen maximiert.

Zur Lösung dieses Problems nutzen wir eine Version eines Satzes von HINDERER [5], den wir an dieser Stelle nur kurz angeben wollen.

Satz 2.5.2
Die Folge der maximalen bedingt erwarteten Gesamtnutzen $\left\{\mathbf{J}_t^*\right\}$ genügt der Optimalitätsgleichung

$$\begin{aligned}\mathbf{J}_t^*(k,b) &= \sup_{a \in A}\left\{n_t(k,b,a) + \beta \int \mathbf{J}_{t+1}\left(k+1, v(b,r), a\right) q_t(k,b;dr)\right\} \\ &=: \mathbf{U}_t \mathbf{J}_{t+1}^*\end{aligned}.$$

Zu einigen Bemerkungen sei auf Anhang A5 verwiesen.

Dieser Satz gibt uns nun die Möglichkeit, unser stochastisches Entscheidungsmodell $M = \left[\left(X, \left\{A(x),\ x \in X\right\}, q_t\left(k,b;\cdot\right), J_t\right)_{t=1,\ldots,T}, X, J_{T+1}\right]$ für die Ausübung Amerikanischer Optionen zu lösen und die Gewinnfunktion iterativ zu bestimmen.

Stellen wir uns nun auf den Standpunkt, daß wir schon am Verfallszeitpunkt der Option angelangt sind, d.h. wir befinden uns im Zeitpunkt t=T. In diesem Zeitpunkt sei unsere Zustandsinformation gegeben durch (k',b'). Setzen wir nun

$$n_t(k',b',a_t) := a_t N(k',b') \ ,$$

dann nimmt die Optimalitätsgleichung folgende Gestalt an:

$$\mathbf{J}_T^*(k,b) = \sup_{a_T \in A}\left\{a_T N(k',b') + (1-a_T)\beta \int \mathbf{J}_{T+1}\left(k'+1, v(b',r)\right) q_T(k',b';dr)\right\} .$$

Diese Gleichung muß nun gelöst werden. Wir erinnern nochmals daran, daß wir uns bei der Wahl der Verteilungen auf die stetigen aus Beispiel 2.4.2 beschränken wollen. Für diese existieren die Erwartungswerte, und sie sind endlich. Damit wird das Supremum auch angenommen und wir können schreiben

$$\mathbf{J}_T^*(k,b) = \max_{a_T \in A}\left\{a_T N(k',b') + (1-a_T)\beta \int \mathbf{J}_{T+1}\left(k'+1, v(b',r)\right) q_T(k',b';dr)\right\} .$$

Wir besitzen jedoch nur zwei Alternativen als Entscheidung. Entweder üben wir die Option aus (a=1) oder wir warten eine weitere Periode, und üben sie somit nicht aus (a=0). Damit nimmt unsere Gleichung folgende Gestalt an;

$$\mathbf{J}_T(k,b) = \begin{cases} N(k',b') & a_T = 1 \\ \beta \int \mathbf{J}_{T+1}\left(k'+1, v(b',r)\right) q_T(k',b';dr) & a_T = 0 \end{cases}$$

Da unsere Option zum Zeitpunkt t=T verfällt, gilt $\mathbf{J}_{T+1} \equiv 0$ und somit

$$\mathbf{J}_T^*(k,b) = \max\left\{N(k',b'), 0\right\} .$$

Ist also der erwartete Barwert der Option größer Null, so werden wir die Option zum Zeitpunkt t=T ausüben, damit lautet unsere Entscheidung für diesen Zeitpunkt

$$a_T^* := \begin{cases} 0 & N(k',b') \le 0 \\ 1 & \text{sonst} \end{cases}$$

Gehen wir einen weiteren Schritt, und nehmen wir an, wir wären erst im Zeitpunkt t=T-1, und wir nehmen weiter an, unsere Zustandsinformation sei (k'',b''). Wir können nun analog wie eben vorgehen. Die Optimalitätsgleichung lautet dann wie folgt:

$$\mathbf{J}_{T-1}^*(k'',b'') = \sup_{a_{T-1}\in A}\left\{a_{T-1}N(k'',b'') + (1-a_{T-1})\beta\int \mathbf{J}_T\left(k''+1, v(b'',r)\right)q_{T-1}(k'',b'';dr)\right\}$$

bzw. aus denselben Gründen wie oben:

$$\mathbf{J}_{T-1}^*(k'',b'') = \max_{a_{T-1}\in A}\left\{a_{T-1}N(k'',b'') + (1-a_{T-1})\beta\int \mathbf{J}_T\left(k''+1, v(b'',r)\right)q_{T-1}(k'',b'';dr)\right\}$$

Um das Maximum zu bestimmen, müssen wir wiederum nur die zwei uns zur Verfügung stehenden Handlungsalternativen betrachten und gelangen zu folgender Gleichung:

$$\mathbf{J}_{T-1}(k'',b'') = \begin{cases} N(k'',b'') & a_{T-1} = 1 \\ \beta\int \mathbf{J}_T\left(k''+1, v(b'',r)\right)q_{T-1}(k'',b'';dr) & a_{T-1} = 0 \end{cases}$$

Damit lautet unsere Optimalitätsgleichung für den Zeitpunkt t=T-1

$$\mathbf{J}_{T-1}^*(k'',b'') = \max\left\{N(k'',b''), \beta\int \mathbf{J}_T(k''+1, v(b'',r))q_{T-1}(k'',b'';dr)\right\}.$$

Damit sind wir wiederum in der Lage, die optimale Entscheidung für den Zeitpunkt t=T-1 anzugeben:

$$a_{T-1}^* = \begin{cases} 0 & N(k'',b'') \le \beta\int \mathbf{J}_T(k''+1, v(b'',r))q_{T-1}(k'',b'';dr) \\ 1 & \text{sonst} \end{cases}$$

Iterativ fahren wir fort bis zum Zeitpunkt t=1. Dort sei unsere Zustandsinformation gegeben durch (k,b). Nach analoger Schlußweise folgt dann für den maximalen bedingt erwarteten Gesamtnutzen

$$\mathbf{J}_1^*(k,b) = \max\left\{N(k,b), \beta\int \mathbf{J}_2(k+1, v(b,r))q_1(k,b;dr)\right\},$$

sowie für die optimale Entscheidung

$$a_1^* = \begin{cases} 0 & N(k,b) \le \beta\int \mathbf{J}_2(k+1, v(b,r))q_1(k,b;dr) \\ 1 & \text{sonst} \end{cases}$$

Wir haben nun das Entscheidungsproblem gelöst, denn mit der Setzung

$$f^* := (a_1^*, \ldots, a_T^*)$$

gemäß

$$a_t^* = \begin{cases} 0: & \text{für } N(k,b) \le \beta\int \mathbf{J}_{t+1}(k+1, v(b,r))q_t(k,b;dr) \\ 1: & \text{sonst} \end{cases} \qquad (2.3.1)$$

für $1 \le t \le T$ stellt f^* die gesuchte optimale Entscheidungsregel für das stochastische Entscheidungsmodell für die Ausübung der Amerikanischen Optionen dar.

Wir haben nun eine recht einfache Entscheidungsregel für unser Ausgangsproblem gefunden, und sie besagt: falls zu einem Zeitpunkt t der erwartete Barwert unserer Option größer ist, als der erwartete verbleibende Gesamtgewinn der folgenden Perioden, so ist es optimal, die Option zum Zeitpunkt t auszuüben.

Da die optimale Entscheidungsregel jedoch eine Funktion der Zustandsinformation (k,b) ist, ist es im allgemeinen nicht optimal, die Option zum Verfallszeitpunkt T auszuüben. Wir werden nun einen Spezialfall angeben, in dem es niemals optimal ist, die Option vor Verfall der Option auszuüben. Dazu betrachten wir einen myopischen (kurzsichtigen) Entscheidungsträger, d. h. einen Entscheidungsträger, der durch einen Zeithorizont der Länge 1 charakterisiert ist. Zum Beginn einer jeden Periode t, $1 \le t \le T-1$, wird er mit zwei Alternativen konfrontiert: entweder übt er die Option zum Zeitpunkt t aus und erhält N(k,b), oder er entscheidet sich, die Option zum Zeitpunkt t+1 die Option auszuüben und $\beta\int N(k+1, v(b,r))q(k,b;dr)$ zu erhalten. Zum Zeitpunkt T übt der Besitzer die Option aus, falls gilt N(k,b)>0.

Satz 2.5.3
Für $(k,b) \in N \times R$ sei

$$N(k,b) \le \beta\int N(k+1, v(b,r))q(k,b;dr)\,, \qquad (2.5.4)$$

dann gilt für $1 \le t \le T-1$:

$$f_t = 0\,.$$

Beweis: Es gilt offensichtlich

$$\beta\int \mathbf{J}_{t+1}\big(k+1, v(b,r)\big)q(k,b;dr) =$$

$$= \beta\int \max\Big\{N(k+1, v(b,r)), \beta\int \mathbf{J}_{t+2}\big(k+2, v(v(b,r), r')\big)\Big\}q(k,b;dr)$$

$$\ge \beta\int N(k+1, v(b,r))q(k,b;dr) \ge N(k,b)$$

Mit (2.5.3) bedeutet dies jedoch nach unserer optimalen Entscheidungsregel $f_t = 0$. •

Der Satz besagt also, falls der myopische Besitzer, der durch (2.5.4) beschrieben wird, die Option nicht ausübt, so gilt dies auch für den Besitzer, der sich optimal verhält.

Es stellt sich nun die Frage, wann denn die Bedingung (2.5.4) erfüllt ist. Dazu nutzen wir ein Resultat aus JAMMERNEGG [6] (siehe Anhang A4), wonach gilt

$$\int N(k+1, v(b,r))q(k,b;dr) = N(k,b)\,.$$

Damit reduziert sich (2.5.4) auf

$$N(k,b) \le \beta N(k,b)\,.$$

Ist also der Diskontfaktor β kleiner oder gleich eins, so ist (2.5.4) erfüllt für **alle** Zustandsinformationen (k,b).

Das heißt: in einem Modell ohne Diskontierung (d. h. β=1) erhalten wir das Resultat von Black und Scholes. Und die Entscheidungsregel für die Ausübung der Amerikanischen Option lautet dann wie folgt:

$$f_t = 0 \quad \forall (k,b) \in N \times R\,,\ 1 \le t \le T-1$$

$$f_T = \begin{cases} 0: & N(k,b) \le 0 \\ 1: & \text{sonst} \end{cases}$$

Es ist somit nicht optimal, die Option vor t=T auszuüben. Und im Zeitpunkt t=T übt der Besitzer die Option nur aus, wenn der erwartete Barwert den Ausübungspreis übersteigt. Ansonsten verfällt die Option unausgeübt.

3. Anhang

A1: Wienerprozeß, Stoppzeiten und Martingale

Bevor wir an dieser Stelle die Begriffe "Wienerprozeß", "Stoppzeit" und "Martingal" einführen können, benötigen wir weitere Begriffe. Wir bezeichnen mit B die Borelsche σ-Algebra.

Definition
Ein stochastischer Prozeß X heißt **meßbar**, wenn für jedes $A \in \mathbf{B}(R^d)$ die Menge $\{(t,\omega); X_t(\omega) \in A\}$ zu der Produkt-σ-Algebra $\mathbf{B}([0,\infty)) \otimes \mathbf{F}$ gehört.

Bemerkung
Wenn wir einen stochastischen Prozeß untersuchen, so setzen wir einen Zeitfluß voraus, d.h. wir können zu jedem Zeitpunkt t≥0 von einer Vergangenheit, Gegenwart und Zukunft sprechen. Das wirft die Frage auf, wieviel wir von unserem Prozeß zu diesem Zeitpunkt wissen, verglichen mit dem, was wir in der Vergangenheit wußten und mit dem, was wir für die Zukunft vermuten. Dazu statten wir unseren Raum $(\Omega,\mathbf{F})$ mit einer **Filtration** aus, d.h. mit einer nichtfallenden Familie $\{\mathbf{F}_t; t \geq 0\}$ von Unter-σ-Algebren von F, für die gilt:

$$\mathbf{F}_s \subseteq \mathbf{F}_t \subseteq \mathbf{F} \text{ mit } 0 \leq s < t < \infty. \text{ Und wir setzen } \mathbf{F}_\infty := \sigma\left(\bigcup_{t\geq 0} \mathbf{F}_t\right).$$

Definition
Ein stochastischer Prozeß X heißt **adaptiert an {F_t}**, wenn für jedes t≥0 X_t eine $\mathbf{F}_t$-meßbare Zufallsvariable ist.

Nun sind wir in der Lage, die wichtigsten der im ersten Kapitel verwendeten Begriffe einzuführen. Wir beginnen mit dem Wienerprozeß.

Definition
Ein **eindimensionaler Standard-Wienerprozeß** ist ein stetiger, adaptierter Prozeß

$$W = \{W_t, \mathbf{F}_t; 0 \leq t < \infty\},$$

der auf dem Wahrscheinlichkeitsraum $(\Omega,\mathbf{F},P)$ definiert ist und folgende Eigenschaften besitzt:

$W_0 = 0$ P-f.s.

$W_t - W_s$ ist für $0 \leq s < t$ unabhängig von $\mathbf{F}_s$ und normalverteilt mit Erwartungswert 0 und Varianz t-s.

Zur Ergänzung wollen wir einige Eigenschaften des Wienerprozesses angeben, die in den Büchern von KARATZAS/SHREVE [8] und ØKSENDAL [12] bewiesen werden.

(i) Der Wienerprozeß ist ein Gauß-Prozeß, d. h. für alle $0 \le t_1 \le ... \le t_k$ besitzt die Zufallsvariable $Z := \left(W_{t_1}, W_{t_2}, ..., W_{t_k}\right) \in R^k$ eine (Multi-) Normalverteilung.

(ii) W besitzt für alle $0 \le t_1 \le ... \le t_k$ unabhängige Zuwächse $W_{t_1}, W_{t_2} - W_{t_1}, W_{t_3} - W_{t_2}, ..., W_{t_k} - W_{t_{k-1}}$

Definition

Es sei ein meßbarer Raum $(\Omega, \mathbf{F})$ gegeben, der mit einer Filtration $\{\mathbf{F}_t\}$ ausgestattet ist. Eine zufällige Zeit T heißt **Stoppzeit** der Filtration, falls das Ereignis $\{T \le t\}$ $\mathbf{F}_t$-meßbar ist für alle t≥0. Eine zufällige Zeit T heißt **Optionszeit**, wenn $\{T \le t\} \in \mathbf{F}_t$ gilt für alle t≥0.

Optional Sampling Theorem

Es sei $\{X_t, \mathbf{F}_t; 0 \le t \le \infty\}$ ein rechts-stetiges Martingal und S≤T zwei Optionszeiten der Filtration $\{\mathbf{F}_t\}$. Dann gilt

$$E\left[X_T \middle| \mathbf{F}_S\right] \ge X_S,$$

sowie

$$EX_T = EX_0.$$

Zum Beweis siehe KARATZAS/SHREVE [8, Satz 1.3.22, S.19].

Wir wenden uns nun dem Begriff des Martingals zu und geben anschließend noch die im Kapitel 1 angeführten Sätze über den Darstellungssatz und die DOOB-MEYER-Zerlegung an.

Definition

Ein Prozeß $\{X_t, \mathbf{F}_t; 0 \le t < \infty\}$ heißt **Submartingal** bzw. **Supermartingal**, wenn für alle $0 \le s < t < \infty$ P-f.s. gilt

$$E\left[X_t \middle| \mathbf{F}_s\right] \ge X_s$$

bzw.

$$E\left[X_t \middle| \mathbf{F}_s\right] \le X_s.$$

Der Prozeß heißt **Martingal**, wenn er sowohl Sub- als auch Supermartingal ist.

Bemerkung

Es seien $\mathbf{S}$ bzw. $\mathbf{S}_a$ die Menge der Stoppzeiten von $\{\mathbf{F}_t\}$ mit

$$P(T < \infty) = 1$$

bzw.

$$P(T \le a) = 1.$$

Dann gehört X=$\{X_t, \mathbf{F}_t; 0 \le t < \infty\}$ zur Klasse D, wenn $\{X_T\}_{T \in \mathbf{S}}$ gleichmäßig integrierbar ist, bzw. zur Klasse DL, falls $\{X_T\}_{T \in \mathbf{S}_a}$ gleichmäßig integrierbar ist für alle 0<a<∞.

DOOB-MEYER-Zerlegung

Es sei $\{\mathbf{F}_t\}$ eine rechtsstetige Filtration und es enthalte $\mathbf{F}_0$ alle P-Null-Ereignisse in

F. Wenn das rechtsstetige Supermartingal $\{X_t,\mathbf{F}_t;0\le t<\infty\}$ aus der Klasse DL ist, dann besitzt es folgende Darstellung:

$$X_t = M_t + A_t,\ 0\le t<\infty,$$

wobei $M=\{M_t,\mathbf{F}_t;0\le t<\infty\}$ ein rechts-stetiges Martingal und $A=\{A_t,\mathbf{F}_t;0\le t<\infty\}$ ein wachsender Prozeß ist.
Diese Zerlegung ist eindeutig.
Ist X aus D, dann ist M gleichmäßig integrierbar und A ist integrierbar.

Zum Beweis siehe KARATZAS/SHREVE [8, Satz 1.4.10, S.24].

Im Beweis von Satz 1.3.5 verwendeten wir eine Version der Bayesschen Formel, die wir an dieser Stelle angeben wollen.

Lemma
Es sei $0\le T<\infty$ fest und es sei Z(X) ein Martingal. Wenn 0≤s≤t≤T gilt und Y eine $\mathbf{F}_t$-meßbare Zufallsvariable ist, für die gilt

$$\tilde{E}_T\big[|Y|\big]<\infty,$$

dann gilt die **Bayessche Formel**:

$$\tilde{E}_T\big[Y|\mathbf{F}_s\big]=\frac{1}{Z_s(X)}E\big[YZ_t(X)|\mathbf{F}_s\big]\ P \text{ und } \tilde{P}_T - \text{f.s.}$$

Auch dazu findet man den Beweis in KARATZAS/SHREVE [8, Lemma 3.5.3, S.193].

Im Kapitel 1 haben wir auch häufig den Darstellungssatz für Martingale verwendet. Dazu geben wir nun zwei Sätze an.

Satz
Es sei $W=\{W_t,\mathbf{F}_t;0\le t<\infty\}$ ein Wienerprozeß auf dem Wahrscheinlichkeitsraum $[\Omega,\mathbf{F},P]$ und es sei $\{\mathbf{F}_t\}$ eine wachsende Folge der durch W erzeugten Filtration.
Dann gilt für ein quadratisch integrierbares Martingal $M=\{M_t,\mathbf{F}_t;0\le t<\infty\}$ mit $M_0=0$ und rechts-stetigen Pfaden mit linksseitigem Grenzwert f.s., daß ein progressiv meßbarer Prozeß $Y=\{Y_t,\mathbf{F}_t;0\le t<\infty\}$ existiert, so daß gilt:

$$E\int_0^T (Y_t)^2\,dt<\infty \text{ für } 0<T<\infty$$

und $$M_t=\int_0^t Y_s\,dW_s \text{ für } 0\le t<\infty.$$

Insbesondere gilt, daß M stetig ist.

Siehe dazu KARATZAS/SHREVE [8, Theorem 3.4.15, S.182].

Satz
Unter den Voraussetzungen des vorhergehenden Satzes sowie, daß ξ eine $\mathbf{F}_\infty$-meßbare Zufallsvariable ist mit $E\xi^2<\infty$, so existiert ein progressiv meßbarer Prozeß Y, derart, daß gilt

$$\xi = E\xi + \int_0^\infty Y_s dW_s \quad \text{P-f.s.}$$

Dieses Resultat wird in KARATZAS/SHREVE [8, Proposition 3.4.18, S.185] bewiesen.

A2: Das Itô-Kalkül und die Itô-Formel

Wir wollen an dieser Selle einen kurzen Einblick in das Itôsche Kalkül geben. Für ein weitergehendes Studium empfehlen wir insbesondere die Lehrbücher von KARATZAS & SHREVE [8] sowie ØKSENDAL [12].

Es sei M die Klasse von Funktionen

$$f(t,\omega):[0,\infty)\times\Omega\to R \quad ,$$

die folgenden Bedingungen genügen:

(i) $(t,\omega)\to f(t,\omega)$ ist $B\times A$ -meßbar, wobei $\mathfrak{B}$ die Borelsche σ-Algebra auf $[0,\infty)$ bezeichnet.

(ii) es existiert eine nichtfallende Familie von σ-Algebren $H_t, t\geq 0$, so daß
  a) W_t ein Martingal bez. $H_t, t\geq 0$ ist und
  b) f_t H_t-adaptiert ist.

(iii) $P\left(\int_0^t f(s,\omega)^2 ds<\infty, \forall t\geq 0\right)=1.$

Definition
Es sei W_t der eindimensionale Standard-Wienerprozeß auf dem Wahrscheinlichkeitsraum $[\Omega, A, P]$. Ein **stochastisches Integral** ist ein stochastischer Prozeß X_t auf $[\Omega, A, P]$ der Form

$$X_t = X_0 + \int_0^t u(s,\omega)ds + \int_0^t v(s,\omega)dW_s \ ,$$

wobei v∈M ist mit

$$P\left(\int_0^t v(s,\omega)^2 ds<\infty, \forall t\geq 0\right)=1.$$

Weiterhin setzen wir voraus, daß u H_t-adaptiert ist und daß gilt

$$P\left(\int_0^t |u(s,\omega)| ds<\infty, \forall t\geq 0\right)=1.$$

Ist X_t ein solches stochastisches Integral, so schreiben wir auch häufig

$$dX_t = u dt + v dW_t \ .$$

Wir kommen nun zum Hauptgegenstand dieses Abschnitts, nämlich der

eindimensionalen Itô-Formel
Es sei X_t ein stochastisches Integral

$$dX_t = u dt + v dW_t \ .$$

Es sei $g(t,x)\in C^2\left([0,\infty)\times R\right)$ (d.h. zweimal stetig differenzierbar). Dann ist

$$Y_t = g(t, X_t)$$

wieder ein stochastisches Integral und es gilt

$$dY_t = \frac{\partial g}{\partial t}(t,X_t)dt + \frac{\partial g}{\partial x}(t,X_t)dX_t + \frac{1}{2}\frac{\partial^2 g}{\partial x^2}(t,X_t)\left(dX_t\right)^2,$$

wobei weiterhin gilt

$$dt \cdot dt = dt \cdot dW_t = dW_t \cdot dt = 0,\ dW_t \cdot dW_t = dt\,.$$

Weiter wollen wir noch die

mehrdimensionalen Itô-Formel angeben.
Es sei

$$dX_t = udt + vdW_t$$

ein n-dimensionales stochastisches Integral. Und sei $g(t,x) = \left(g_1(t,x),...,g_p(t,x)\right)$ eine C^2 Abbildung von $\left[0,\infty\right) \times R^n$ in R^p, dann ist der Prozeß

$$Y(t,\omega) = g(t,X_t)$$

wieder ein stochastisches Integral, dessen Komponenten gegeben sind durch

$$dY_k = \frac{\partial g_k}{\partial t}(t,X)dt + \sum_i \frac{\partial g_k}{\partial x_i}(t,X)dX_i + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g_k}{\partial x_i \partial x_j}(t,X)dX_i dX_j\,,$$

wobei weiterhin gilt

$$dt \cdot dt = dt \cdot dW_i = dW_i \cdot dt = 0,\ dW_i \cdot dW_j = \delta_{ij}dt\,.$$

Den Beweis und viele Beispiele dazu findet man in ØKSENDAL [12], sowie in KARATZAS/SHREVE [8].

A3: Die Novikov-Bedingung und das Girsanov-Theorem

Novikov-Bedingung

Es sei $W = \{W_t, F_t;\ 0 \le t < \infty\}$ ein eindimensionaler Standard-Wienerprozeß und $Y = \{Y_t,\ F_t;\ 0 \le t < \infty\}$ ein meßbarer, adaptierter Prozeß, der folgender Bedingung genügt:

$$P\left(\int_0^T Y_t^2 dt < \infty\right) = 1$$

und wenn gilt:

$$E\left[\exp\left\{\frac{1}{2}\int_0^T \|Y_s\|^2 ds\right\}\right] < \infty,$$

so ist $Z = e^Y$ ein Martingal.

Einen Beweis findet man dazu in KARATZAS/SHREVE [8, Seite 199].

Girsanov-Theorem

Es sei ein Martingal $Z_t = \exp\left\{\int_0^t \theta_s dW_s - \frac{1}{2}\int_0^t \|\theta_s\|^2 ds\right\}$ gegeben. Wird nun $\widetilde{W} = \{\widetilde{W}_t,\ F_t;\ 0 \le t < \infty\}$ definiert gemäß

$$\widetilde{W}_t := W_t - \int_0^t \theta_s ds, \quad 0 \le t < \infty,$$

dann ist $\widetilde{W}$ ein eindimensionaler Wienerprozeß auf $[\Omega,\ F_T,\ \widetilde{P}_T]$, $\forall T \in [0,\infty)$ und mit $\widetilde{P}_T(A) := E[Z_T 1_A], \quad A \in F_T$.

Auch dieser Beweis kann in KARATZAS/SHREVE [8, Kapitel 3.5., Seite 190ff] nachgelesen werden.

A4: Bedingte Erwartung, Stichprobe, Statistik

In diesem Teil des Anhangs wollen wir kurz den bedingten Erwartungswert einführen und einige Eigenschaften aufzählen. Für ein weiteres Studium wollen wir an dieser Stelle zum Beispiel auf ØKSENDAL [12] verweisen.

Es sei $[\Omega, \mathbf{A}, P]$ ein Wahrscheinlichkeitsraum und es sei $X: \Omega \to R^n$ eine Zufallsvariable, für die gilt: $E[|X|] < \infty$.

Definition
Es sei $\mathbf{H} \subset \mathbf{A}$ eine σ-Algebra. Die f.s.-eindeutige Funktion $E[X|\mathbf{H}]: \Omega \to R^n$ heißt **bedingte Erwartung von X unter H**, wenn sie folgenden Bedingungen genügt:

(i) $E[X|\mathbf{H}]$ ist H-meßbar

(ii) $\int_H E[X|\mathbf{H}]dP = \int_H XdP, \ \forall H \in \mathbf{H}$

Es sei $Y: \Omega \to R^n$ eine weitere Zufallsvariable mit $E[|Y|] < \infty$, und es seien a,b zwei reelle Zahlen. Dann gelten folgende Eigenschaften:

I) $E[aX + bY|\mathbf{H}] = aE[X|\mathbf{H}] + bE[Y|\mathbf{H}]$,

II) $E[E[X|\mathbf{H}]] = E[X]$,

III) $E[X|\mathbf{H}] = X$, falls X H-meßbar ist und

IV) $E[X|\mathbf{H}] = E[X]$, falls X unabhängig ist von H.

Die Beweise dieser Eigenschaften sind recht einfach und werden an dieser Stelle nicht geführt, können jedoch in oben angegebener Literatur nachgeschlagen werden.

Im zweiten Kapitel spielten die Begriffe "Stichprobe" und "Statistik" eine wichtige Rolle. Auch diese Begriffe wollen wir anschließend kurz einführen.

Definition
Auf dem Wahrscheinlichkeitsraum $[\Omega, \mathbf{A}, P]$ seien unabhängige Zufallsgrößen $X_1, X_2, ..., X_n$ mit derselben Verteilungsfunktion F definiert

$$F(x) = P(X_k \leq x), \ \ k = 1,2,...,n.$$

Wir nennen $Z := (X_1, X_2, ..., X_n)$ **(mathematische) Stichprobe vom Umfang n** aus einer nach F verteilten Grundgesamtheit.

Für $\omega_0 \in \Omega$ heißt $Z(\omega_0) = (x_1, x_2, ..., x_n)$ dann **konkrete Stichprobe** oder Realisierung der Stichprobe $(X_1, X_2, ..., X_n)$. Wegen der Unabhängigkeit der Elemente und der identischen Verteilungen ist die Verteilungsfunktion F_Z der Stichprobe Z gegeben durch

$$F_Z(x_1, ..., x_n) := \prod_{k=1}^{n} F(x_k).$$

Eine meßbare Funktion der Stichprobenvariablen $X_1, X_2, ..., X_n$ bezeichnet man als **Statistik** oder auch als **Stichprobenfunktion**. Die Statistik selbst ist somit eine zufällige Variable. Ihre Realisierung, d.h. die entsprechende Funktion der konkreten Stichprobe, heißt konkrete Statistik.

Wir benutzten im Abschnitt 2.5. ein Resultat aus JAMMERNEGG [6, Satz 2.5, S.15], das wir jetzt kurz einführen wollen.

Die Wahrscheinlichkeits- bzw. Dichtefunktion der suffizienten Statistik Y wird mit $g(n, y; \vartheta)$ bezeichnet, wobei $\vartheta \in \Theta$ der unbekannte Parameter und (n,y) die gegebene Zustandsinformation sind. Ausgehend von der A-priori-Wahrscheinlichkeitsfunktion bzw. -dichte $\overline{\mu}$ für den unbekannten Parameter $\vartheta \in \Theta$ berechnet sich dann die A-posteriori-Wahrscheinlichkeitsfunktion bzw. -dichte nach der Bayesschen Formel für Dichten wie folgt:

$$\mu(n, y; D) := \frac{\int_D g(n, y; \vartheta)\overline{\mu}(d\vartheta)}{\int_\Theta g(n, y; \vartheta')\overline{\mu}(d\vartheta')}, \text{ für } n=1,2,...$$

und $$\mu(n, y; D) := \overline{\mu}(D), \qquad \text{für } n=0,$$

mit $D \subseteq \Theta$. Wir sind nun in der Lage, die A-posteriori-Verteilung p für die Zufallsvariable X_{n+1} anzugeben, d.h. in unserem Zusammenhang die Verteilung des zufälligen Marktwertes in der nächsten Periode unter der Kenntnis der Zustandsinformation (n,y):

$$p(n, y; B) := \int_\Theta \nu_\vartheta(B)\mu(n, y; d\vartheta).$$

Diese Verteilung ist wieder eine Übergangswahrscheinlichkeit. Für alle Funktionen $U: X \to R$ wird die bedingte Erwartung bei gegebenen (n,y) berechnet gemäß

$$u(n, y) := \int_X U(x)p(n, y; dx).$$

Weiter wissen wir aus Beispiel 2.2.5, daß für die dort angeführten Verteilungen die suffiziente Statistik $Y = T(X_1, ..., X_n)$ auf folgende Weise rekursiv bestimmt werden kann:

$$T(X_1, ..., X_{n+1}) = v\big(T(X_1, ..., X_n), X_{n+1}\big) = v(Y, X_{n+1}),$$

wobei die v(y,x) gegeben ist entweder durch v(y,x)=y+x oder v(y,x)=yx (für die Log-Normalverteilung).

Satz
Für alle Zustandsinformationen $(k,b) \in N_0 \times R$ gilt

$$\int u(k+1, v(b,r))q(k,b;dr) = u(k,b)$$

Wenn wir U als Nutzenfunktion interpretieren, dann besagt die Behauptung des Satzes: wenn wir von der Zustandsinformation (n,y) ausgehen, so ist der erwartete Nutzen für n+1 Beobachtungen genauso groß wie der für n Beobachtungen berechnete.

A5: Zur dynamischen Optimierung

Im Abschnitt 2.5 sind wir nach dem Satz 2.5.2 von HINDERER noch einige Bemerkungen schuldig geblieben.

Zuvor müssen wir jedoch unser Entscheidungsmodell ein wenig präzisieren und führen daher weitere Begriffe ein. Wir betrachten weiterhin unser Modell

$$M = \left[\left(X, \{0,1\}, q_t(k,b;\cdot), \mathbf{J}_t\right)_{t=1,\dots,T}, X, \mathbf{J}_{T+1}\right].$$

- Der Zustandsraum $\mathfrak{X}$ unseres Entscheidungsmodells sei ausgestattet mit einer σ-Algebra γ.
- Weiter sei der Aktionenraum $\mathfrak{A}$={0,1} ausgestattet mit der σ-Algebra δ.
- Es sei $H_t := X \times A \times \cdots \times X$ (2n-1)-mal die Vergangenheit unseres Prozesses mit den Elementen $h_t = (x_1, a_1, \dots, x_t)$. Die Produkt-σ-Algebra von H_t sei η_t. Es sei D die Folge von Abbildungen D_t von $H_t \subset H_t$ in die Menge aller nichtleeren Untermengen von $\mathfrak{A}$ mit $H_1 := X$ und der Eigenschaft, daß $H_{t+1} := \left\{(h,a,x) \in H_{t+1} : h \in H_t, a \in D_t(h), x \in X\right\}$ zu η_{t+1} gehört für alle t>0. Weiter sei $D_t(h)$ die Menge aller zulässigen Aktionen zum Zeitpunkt t und bei Vergangenheit h. Die Übergangswahrscheinlichkeit von H_t nach $\mathfrak{A}$ sei gegeben durch π_t mit $\pi_t(h, D_t(h)) = 1$. Wir bezeichnen $\pi := (\pi_t)_{t>1}$ als zufällige Entscheidungsregel und Δ als Menge aller zufälligen Entscheidungsregeln.
- Der bedingt erwartete Nutzen bei der Entscheidungsregel α und der Zustandsinformation (k,b) sei $\mathbf{J}_{t,\alpha} := \mathbf{J}_{t,\alpha}(k,b)$. Der **maximale bedingt erwartete Nutzen** wird dann definiert gemäß $\mathbf{J}_t(h) := \sup_{\alpha \in \Delta}\left\{\mathbf{J}_{t,\alpha}(h)\right\}$.

Satz
Die Folge $\{\mathbf{J}_t\}_{t=1,\dots,T}$ genügt der Optimalitätsgleichung

$$\mathbf{J}_t = \sup_{a \in D_t(\cdot)}\left\{n_t(\cdot,a) + \int \mathbf{J}_{t+1}(\cdot,a,x) q_t(\cdot,a,dx)\right\}$$
$$=: \mathbf{U}_t \mathbf{J}_{t+1}$$

Der Beweis kann mit Hilfe der oben genannten Begriffe bei HINDERER [5, Theorem 14.4., S.101] nachgelesen werden.

Dieser Satz liefert uns die Optimalität für eine Folge von Nutzenfunktionen. Diese Nutzenfunktionen sind jedoch Funktionen von der Vergangenheit H_t unseres Prozesses. In dem Modell, das wir im Abschnitt 2.5 entwickelten, sind diese Nutzenfunktionen aber von einer suffizienten Statistik b abhängig. Dies führt uns nun zu folgender Definition:

Definition
Es sei $b = (b_k)$ eine suffiziente Statistik bez. des unbekannten Parameters ϑ für unser Entscheidungsmodell. Dann heißt eine zufällige Entscheidungsregel $\pi \in \Delta$ eine **b-Entscheidungsregel**, falls Übergangswahrscheinlichkeiten π'_n existieren, so daß gilt:

$$\pi_n(h,\eta) = \pi'_n(b(h),\eta) \quad n \in N, h \in H_n, \eta \in \eta_n .$$

Wir führen nun noch einige weitere Begriffe ein, die es uns gestatten, schließlich die Optimalität unserer Gewinnfunktionen in einem Satz zu formulieren.

Es sei Q_n die Menge der Abbildungen $w: b(H,n) \to R$, derart daß $w \circ b(n)$ von unten beschränkt und meßbar ist. Wir definieren nun noch den Operator $\mathbf{U}'_n$ gemäß

$$\mathbf{U}'_k w := \sup_{a \in D'_k(b(h))} \left\{ n'_k(b(h),a) + \int w(x) q'_k(b(h),a,dx) \right\},\ h \in H_k \text{ und } \forall k \in N .$$

Definition
Es sei b(h) $h \in H_k$ eine suffiziente Statistik bez. des unbekannten Parameters ϑ unseres Entscheidungsproblems. Eine Lösung der **reduzierten Optimalitätsgleichung** ist eine Folge (w_k) von Abbildungen $w_k \in Q_k$, so daß gilt:

$$w_k = \mathbf{U}'_k w_{k+1},\ k \in N ,$$

bzw.
$$w_k := \sup_{a \in D'_k(b(h))} \left\{ n'_k(b(h),a) + \int w_{k+1}(x) q'_k(b(h),a,dx) \right\},\ h \in H_k \text{ und } \forall k \in N .$$

Nun können wir folgenden Satz abgeben:

Satz
Es sei $b = (b_k)$ eine suffiziente Statistik bez. des unbekannten Parameters ϑ. Weiter sei Δ' die Menge aller b-Entscheidungsregeln. Dann gilt

i) $\mathbf{J}_k = \sup_{\pi' \in \Delta'} \mathbf{J}_{\pi',k}$

ii) Es existiert für alle $k \in N$ eine eindeutige Abbildung $\mathbf{J}'_k : b_k(H) \to R$ mit $\mathbf{J}_k = \mathbf{J}'_k \circ b_k$

iii) Die Folge $(\mathbf{J}'_k)$ ist Lösung der reduzierten Optimalitätsgleichung.

Einen Beweis für diesen Satz findet man in HINDERER [5, Theorem 18.4, S. 122].

Wir können also unser Entscheidungsproblem mit Hilfe der suffizienten Statistiken, die wir im Kapitel 2 einführten, lösen.

4. Kurzzusammenfassung

Wir haben uns in der vorliegenden Arbeit mit Amerikanischen Optionen beschäftigt. Im ersten Kapitel entwickelten wir dazu ein Marktmodell, in dem nur zwei Wertpapiere vorhanden sind: der Bond als festverzinsliches Wertpapier und die risikobehaftete Anlage, wie zum Beispiel eine Aktie.

Wir haben weiter angenommen, daß die Zinsrate, mit dem der Bond verzinst wird, konstant ist, und auch der Streuungskoeffizient des Preisprozesses für unser risikobehaftetes Wertpapier sollte konstant sein. Dadurch war es uns möglich, die Black und Scholes-Formel in expliziter Form herzuleiten. Sobald man jedoch die Konstantheit dieser beiden Marktkoeffizienten aufgibt, ist dies nicht mehr ohne weiteres möglich.

Im ersten Kapitel war es unser Hauptziel, einen fairen Preis für die Amerikanischen Optionen zu bestimmen. Um dies zu bewerkstelligen, bedienten wir uns der Martingaltheorie und insbesondere des Darstellungssatzes für Martingale, wie er in KARATZAS/SHREVE [8] gegeben wird. Zusammen mit der Itô-Formel war es uns möglich, den fairen Preis zu bestimmen. Weiterhin haben wir in den Beweisen der Sätze 1.3.5 und 1.5.3 den Portefeuilleprozeß bestimmen können, der den Hedging-Prozeß gegen die Option darstellt, mit dessen Hilfe der Preis bestimmt werden konnte.

Im zweiten Kapitel beschäftigten wir uns mit einem diskreten Modell für die Bestimmung der optimalen Entscheidungsregel für die Ausübung der Amerikanischen Option. Die tragende theoretische Säule war hierbei der Bayessche Satz für Dichten, der auch in der Arbeit bewiesen wurde. Mit diesem Satz war es uns dann möglich, unser stochastisches Entscheidungsmodell aufzustellen und das dazugehörige Entscheidungsproblem zu formulieren und zu lösen.

Wir haben in dieser Arbeit jedoch nur einen kleinen Teil der Problematik um Amerikanische Optionen beleuchten können. Für das weitere Studium ist zur Einstimmung auf jeden Fall die Arbeit von KARATZAS [7] mit der dort angegebenen Literatur, sowie die Dissertation von KORN [9] zu empfehlen. Im Hinblick auf das diskrete Modell im zweiten Kapitel ist zu erwähnen, daß JAMMERNEGG [6] das Problem für eine größere Klasse von Investitionsentscheidungsproblemen löst.

5. Literaturverzeichnis

[1] BERGER,JO: Statistical Decision Theory. Springer Verlag 1980.

[2] BISMUT,JM / SKALLI,I: Temps d'arrêt optimal, théorie générale des processus et processus de Markov. Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete 39: 301-313. Springer 1977.

[3] BLACK,F / SCHOLES,M.: The Pricing of Options and Corporate Liabilities. Journal of Political Economics 81: 637-653. 1973.

[4] GIRLICH,H.-J.: Diskrete stochastische Entscheidungsprozesse. Teubner 1973.

[5] HINDERER,K.: Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter. Lecture Notes in Operations Research and Mathematical Systems 33. Springer 1970.

[6] JAMMERNEGG,W: Sequential Binary Investment Decisions. Lecture Notes in Economics and Mathematical Systems 313. Springer Verlag 1988.

[7] KARATZAS, I.: On the Pricing of American Options. Applied Mathematics and Optimization 17: 37-60, 1988.

[8] KARATZAS, I. / SHREVE, S.E.: Brownian Motion and Stochastic Calculus. Springer-Verlag 1991.

[9] KORN,R.: Optimierungsprobleme bei Wertpapierhandel in stetiger Zeit. Dissertation im Fachbereich Mathematik der Johannes Gutenberg-Universität Mainz. Mainz 1992.

[10] MERTON,RC: Theory of rational option pricing. Bell Journal of Economic Management Sciences 4: 141-183, 1973.

[11] MÜLLER, P.H. (Herausgeber): Wahrscheinlichkeitsrechnung und Mathematische Statistik. Lexikon der Stochastik. Akademie Verlag, 1991.

[12] ØKSENDAL, B.: Stochastic Differential Equations. Springer-Verlag 1992.

[13] RASCH,D.: Einführung in die mathematische Statistik. Berlin 1989.

[14] SPREMANN, K.: Investition und Finanzierung. Oldenbourg Verlag, 1990.

[15] VIERTL: Einführung in die Stochastik. Springer Verlag 1990.

Erklärung

Ich versichere, daß ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Quellen und Hilfsmittel angefertigt habe.

Leipzig (Datum) (Unterschrift)