Optimal Selection from Distributions with Unknown Parameters: Robustness of Bayesian Models.

Alfred Müller

Institut für Wirtschaftstheorie und Operations Research, Universität Karlsruhe, Kaiserstr. 12, D-76128 Karlsruhe, Germany

Abstract

We consider the problem of making one choice from a known number of i.i.d. alternatives. It is assumed that the distribution of the alternatives has some unknown parameter. We follow a Bayesian approach to maximize the discounted expected value of the chosen alternative minus the costs for the observations. For the case of gamma and normal distribution we investigate the sensitivity of the solution with respect to the prior distributions. Our main objective is to derive monotonicity and continuity results for the dependence on parameters of the prior distributions. Thus we prove some sort of Bayesian robustness of the model.

Keywords: Bayesian Dynamic Programming, Stochastic Orders, Kantorovich metric, Optimal Selection, Sensitivity Analysis, Monotonicity, Bayesian Robustness.

1 Introduction

There are many situations, where one has to select one offer of sequentially arriving alternatives. Special examples that have been considered in the literature are the secretary problem, the problem of job search or the problem of selling an asset. There are several possibilities for modelling this situation.

In this paper we investigate the following model: There is a fixed number N of alternatives. The decision maker has to choose exactly one of these offers. The alternatives can be described as independent and identically distributed real random variables $X_1, ..., X_N$, which can be observed sequentially at a cost c per observation. We consider two possibilities of the reward structure. If the decision maker stops after the nth observation, $1 \le n \le N$, his reward is either X_n, if no recall is allowed, or $\max\{X_1, ..., X_n\}$, if recall is allowed. We assume that the common distribution of the observations involves a parameter θ, which is unknown to the decision maker. We follow a Bayesian approach to find a strategy that maximizes the discounted expected total reward.

This problem has been considered in the literature for several choices of distributions. Sakaguchi (1961) and DeGroot (1968) regarded this problem under the assumption of a normal distribution with unknown mean. Stewart (1978) obtained some results for the case of a uniform distribution with unknown endpoints and Tamaki (1983) studied it for observations drawn from a gamma distribution.

The main objective of this paper is to investigate the sensitivity of the solution with respect to parameters of the prior distribution. This is an important topic for justifying the use of the Bayesian approach, since there is always some uncertainty in the elicitation of a prior distribution. Using appropriate notions of stochastic order relations and probability metrics, we derive (for the model with normal and gamma distributed alternatives) monotonicity and continuity results for the dependence on parameters of the prior distribution. As a by-product of this approach we prove in Theorem 4.2 a conjecture of Tamaki (1983).

Our paper is organized as follows. In section 2 we define a Bayesian control model, that gives a formal description of the problem mentioned above. Section 3 contains some well known facts about stochastic orders and probability metrics. These are used in section 4 to derive the main results about monotonicity and continuity. In section 5 we compare our analytical bounds with computational results.

2 The Bayesian Control Model.

The problem described in the introduction can be modelled as a Bayesian Control Model (BCM) as introduced in Rieder (1975, 1988), see also Rieder and Wagner (1991).
The BCM is given by a tupel $(S, A, D, Z, T, \Theta, \mu_0, Q, r, V_0, \beta)$ which in our case has the following meaning:

(i) The **state space** $S := M \cup \{\infty\}$ with the following interpretation: $M \subset \mathbb{R}$ is the set of possible alternatives and $s \in M$ is the best momentary available alternative. If you have already accepted an offer, then you are in state $s = \infty$.

(ii) The **action space** $A := \{0, 1\}$ has the following meaning: If you accept an alternative then $a = 1$, and if you reject it then $a = 0$.

(iii) The **restriction set** $D \subset S \times A$ desribes, which actions a are allowed in state s. In our case we have $D = S \times A/\{\infty, 1\}$. This means, that if we have accepted an offer ($s = \infty$) then we must reject all subsequent offers ($a \neq 1$).

(iv) The **disturbance space** $Z := M$, as in our case the "disturbances" are the new offers.

(v) The **transition function** $T : D \times Z \to S$ describes the transition from the momentary state s to the new state s' under action a and disturbance z. If we accept an alternative then the new state is $s' = \infty$. Hence $T(s, 1, z) = \infty$ for all s and z. If we reject an offer, then the new state is the next offer z if no recall is allowed and $\max\{s, z\}$ if recall is allowed. Thus we have $T(s, 0, z) = z$ resp. $T(s, 0, z) = \max\{s, z\}$ for all z and all $s \in M$. Sometimes we will use the convenient notion $s \vee z := \max\{s, z\}$.

(vi) $\Theta \neq \emptyset$ is an arbitrary **parameter space** for the unknown parameter of the distribution $Q(\theta, dz)$, $\theta \in \Theta$, of the offers. We assume that there is a σ-finite measure ν on Z, such that $Q(\theta, \cdot)$ has a density $q(\theta, \cdot)$ with respect to ν for all $\theta \in \Theta$.

(vii) The **prior distribution** μ_0 is an arbitrary probability measure on Θ, which represents the initial knowledge about the unknown parameter.

(viii) The **one step reward function** $r : D \to \mathbb{R}$ describes the reward that we get in one period. If we accept an offer then we get the value s of the best actually available offer; if we reject it, we have to pay the observation cost c; and if we are already in the absorbing state $s = \infty$ then we get nothing. Hence we have

$$r(s,a) := \begin{cases} 0, & s = \infty \\ s, & s \in M,\ a = 1 \\ -c, & s \in M,\ a = 0 \end{cases} .$$

(ix) The **terminal reward function** $V_0 : S \to \mathbb{R}$ determines the reward, that we get at the end of the procedure after the N-th period. It is equal to s, if $s \in M$ and 0, if $s = \infty$.

(x) Finally, all rewards are discounted in each period by a **discount factor** $\beta \in (0,1]$.

For each parameter $\theta \in \Theta$, we denote by $V^{\theta}_{N\pi}(s)$ the expected discounted reward, when we start is state $s \in S$, apply the (history dependent) N-stage policy π, and if $Q(\theta,\cdot)$ is the distribution of the alternatives.

Following the Bayesian approach, our aim is to maximize the **Bayesian expected reward**

$$V_{N\pi}(s) := \int \mu_0(d\theta)\ V^{\theta}_{N\pi}(s)$$

over all N-stage policies π.

It is well known, that the Bayesian Control Model can be reduced to a Markovian Control Model (MCM) by augmentation of the state space to $\hat{S} := S \times \mathbb{P}(\Theta)$, where $\mathbb{P}(\Theta)$ is the set of all probability measures on Θ. Moreover, if there is a so called *sufficient statistic*, then we can diminish $\mathbb{P}(\Theta)$ to an *information space* I. A formal description of this procedure can be found in Hinderer (1970) and Rieder (1975, 1988). It is skipped here.

In our model the value iteration then assumes the following form:

$$V_k(i,s) = \max\{s, c_k(i,s)\},\ k \in \mathbb{N},\ i \in I,\ s \in M, \tag{2.1}$$

where

$$c_k(i,s) := -c + \beta \int Q(i,dz)\ V_{k-1}(\phi(i,z), T(s,a,z)). \tag{2.2}$$

Here the function ϕ describes the bayesian updating of the parameter i under the observation z, i.e. if $\mu(i,\cdot)$ is the prior distribution and we observe z, then $\mu(\phi(i,z),\cdot)$ is the corresponding posterior distribution. The transition probability measure $Q(i,\cdot)$ is defined as

$$Q(i,dz) := \int \mu(i,d\theta)\ Q(\theta,dz). \tag{2.3}$$

The state $s = \infty$ is an absorbing state with $V_k(i, \infty) = 0$ for all $k \in \mathbb{N}$ and $i \in I$. Therefore, this state will not be considered any more.

Example 2.1. (cf. DeGroot (1968))
Assume that $Q(\theta, dz) = \mathcal{N}(\theta, \nu)$, a normal distribution with unknown mean θ and known variance ν. Assume further that the prior distribution is also a normal distribution $\mu_0 = \mathcal{N}(x, b)$. Then it is well known, that after observing z, one gets as posterior distribution again a normal distribution

$$\mu_1 = \mathcal{N}(\frac{x\nu + zb}{\nu + b}, \frac{\nu b}{\nu + b}).$$

Hence it is sufficient to know the value of the mean and the variance of the current posterior distribution. Thus we can choose $I = \mathbb{R} \times \mathbb{R}_{>0}$ as information space. From (2.3) we deduce that $Q(x, b, \cdot) = \mathcal{N}(x, b + \nu)$. The value iteration then assumes the form

$$V_{k+1}(x, b, s) = \max\left\{s, -c + \beta \int \mathcal{N}(x, b + \nu, dz)\ V_k\left(\frac{x\nu + zb}{\nu + b}, \frac{\nu b}{\nu + b}, T(s, a, z)\right)\right\} \tag{2.4}$$

$\square$

Example 2.2. (cf. Tamaki (1983))
Let $Q(\theta, dz) = \,,(\theta, \lambda)$ be a gamma distribution with unknown scale parameter θ and known shape parameter λ. Assume that the prior distribution is also a gamma distribution, $\mu_0 = \,,(x, y)$. Then one gets as posterior distribution again a gamma distribution $\mu_1 = \,,(x + z, y + \lambda)$. Hence one can choose $I = \mathbb{R}_{>0} \times \mathbb{R}_{>0}$. From (2.3) we get $Q(x, y, \cdot) = \mathrm{GB2}(\lambda, y, x)$, where $\mathrm{GB2}(\alpha, \delta, \gamma)$ is the *generalized beta distribution of the second kind* with the density

$$z \;\to\; \frac{,(\alpha + \delta)}{,(\alpha) \cdot ,(\delta)} \cdot \frac{z^{\alpha - 1}\gamma^{\delta}}{(\gamma + z)^{\alpha + \delta}}, \quad \alpha, \delta, \gamma, z \in \mathbb{R}_{>0}.$$

This distribution has been considered e.g. by Cummins et.al. (1990).

The value iteration is then given by

$$V_{k+1}(x, y, s) = \max\{s, -c + \beta \int \mathrm{GB2}(\lambda, y, x, dz)\ V_k(x + z, y + \lambda, T(s, a, z)). \tag{2.5}$$

$\square$

If no recall is allowed, then the function c_k defined in (2.2) is independent of s and hence it is optimal at stage k to accept an offer s iff

$$s \geq c_k(i) := -c + \beta \int Q(i, dz)\ V_{k-1}(\phi(i, z), z).$$

Thus we have an optimal policy of control limit type. We will prove now that this is also true if recall is allowed.

Theorem 2.1 *Assume that recall is allowed. Then it is optimal to accept the offer s at stage k iff $s \geq c_k'(i)$, where*

$$c_k'(i) := \inf\{s : s \geq c_k(i, s) = -c + \beta \int Q(i, dz)\ V_{k-1}(\phi(i, z), s \vee z)\}.$$

Proof. This follows immediately from (2.1), if we can show that $s \to s - c_k(i,s)$ is non-decreasing. But this follows easily by induction on k. $\square$

3 Stochastic Orders and Probability Metrics.

The main purpose of this paper is to show monotonicity and continuity results of the functions $i \to V_k(i)$ resp. $i \to c_k(i)$. For this we need appropriate concepts of stochastic orders and probability metrics.

Definition 3.1 *a) Let P, Q be probability measures with distribution functions F resp. G. Then P is said to be stochastically smaller than Q (written $P \leq_{st} Q$), if $F(x) \geq G(x)$ for all $x \in \mathbb{R}$.*

b) If P, Q have densities f resp. g with respect to some σ-finite measure ρ, such that

$$f(y)g(x) \leq f(x)g(y) \quad \textit{for all} \quad x \leq y,$$

then P is said to be smaller than Q in likelihood ratio order (written $P \leq_{lr} Q$).

A concise treatment of these orderings can be found in Shaked and Shanthikumar (1994). The most important facts that we need here are summarized in the following theorem.

Theorem 3.2 *a) The relation $P \leq_{st} Q$ holds iff $\int f \, dP \leq \int f \, dQ$ for all increasing functions, such that the integrals exist.*

b) $P \leq_{lr} Q$ implies $P \leq_{st} Q$.

Example 3.1. a) For normal distributions we have the following result:
$\mathcal{N}(\mu_1, \sigma_1^2) \leq_{lr} \mathcal{N}(\mu_2, \sigma_2^2)$ iff $\mathcal{N}(\mu_1, \sigma_1^2) \leq_{st} \mathcal{N}(\mu_2, \sigma_2^2)$ iff $\mu_1 \leq \mu_2$ and $\sigma_1^2 = \sigma_2^2$.

b) For the case of gamma distributions we get
$, (\alpha_1, \nu_1) \leq_{lr} , (\alpha_2, \nu_2)$ iff $, (\alpha_1, \nu_1) \leq_{st} , (\alpha_2, \nu_2)$ iff $(\alpha_1 \geq \alpha_2$ and $\nu_1 \leq \nu_2)$.

c) The family of generalized beta distributions GB2(α, δ, γ) introduced in Example 2.2 is $\leq_{lr}$-increasing in α and γ and $\leq_{lr}$-decreasing in δ. Due to Theorem 3.2 b) the same is true for the ordering $\leq_{st}$. $\square$

Next we want to introduce a probability metric that is valuable for proving continuity results. It turns out, that the Kantorovich metric is a suitable concept.

For a Lipschitz function $f : \mathbb{R} \to \mathbb{R}$ we define the Lipschitz seminorm

$$\|f\|_L \ := \ \sup_{x \neq y} \frac{|f(x) - f(y)|}{|x - y|}.$$

We denote by $\mathfrak{L}$ the set of all Lipschitz functions and $\mathfrak{L}_1$ shall be the set of all Lipschitz functions f with $\|f\|_L \leq 1$.

$\mathbb{P}_m$ will denote the collection of all probability measures with finite mean. Then the Kantorovich metric is defined as follows, cf. Zolotarev (1983) or Rachev (1991).

Definition 3.3 *For any two probability measures $P, Q \in \mathbb{P}_m$ we define the* Kantorovich distance

$$\zeta(P,Q) \ := \ \sup \left\{ \left| \int f \ dP - \int f \ dQ \right| : f \in \mathfrak{L}_1 \right\}.$$

For random variables X, Y with distributions $P_X, P_Y \in \mathbb{P}_m$ we write $\zeta(X,Y) := \zeta(P_X, P_Y)$.

It is well known that the Kantorovich metric is equivalent to the L^1-distance of the corresponding cumulative distribution functions, i.e. if F and G are the c.d.f.'s of P resp. Q, then

$$\zeta(P,Q) \ = \ \int_{-\infty}^{\infty} |F(x) - G(x)| \ dx, \tag{3.1}$$

cf. Rachev (1991), p. 6. Using (3.1) it is very often possible to calculate the Kantorovich distance explicitly. Some important cases are treated in the following theorem. The easy proof is omitted.

Theorem 3.4 *a) If X, Y are integrable random variables with $X \leq_{st} Y$, then*

$$\zeta(X,Y) \ = \ EY - EX.$$

b) Assume X is an integrable random variable and $Y := aX$ with $a > 1$. Then $\zeta(X,Y) = (a-1) \cdot E|X|$.

c) $\zeta(aX, aY) = a \cdot \zeta(X,Y)$ for all $a > 0$.

d) $\zeta(X+a, Y+a) = \zeta(X,Y)$ for all $a \in \mathbb{R}$.

Example 3.2. From Theorem 3.4 the following examples can easily be derived.

a) If $X \sim \mathcal{N}(x,b)$ and $Y \sim \mathcal{N}(y,b)$, then $\zeta(X,Y) = |x-y|$.

b) If $X \sim \mathcal{N}(x,b_1)$ and $Y \sim \mathcal{N}(x,b_2)$, then $\zeta(X,Y) = \sqrt{2/\pi} \cdot |\sqrt{b_1} - \sqrt{b_2}|$.

c1) If $X \sim \mathrm{GB2}(\alpha_1, \delta, \gamma)$, $Y \sim \mathrm{GB2}(\alpha_2, \delta, \gamma)$ and $\delta > 1$, then

$$\zeta(X,Y) = |\alpha_1 - \alpha_2| \cdot \gamma / (\delta - 1).$$

c2) If $X \sim \mathrm{GB2}(\alpha, \delta, \gamma_1)$, $Y \sim \mathrm{GB2}(\alpha, \delta, \gamma_2)$ and $\delta > 1$, then

$$\zeta(X,Y) = |\gamma_1 - \gamma_2| \cdot \alpha / (\delta - 1).$$

c3) If $X \sim \mathrm{GB2}(\alpha, \delta_1, \gamma)$ and $Y \sim \mathrm{GB2}(\alpha, \delta_2, \gamma)$ and $\delta_1, \delta_2 > 1$ then

$$\zeta(X,Y) = \alpha\gamma \cdot \left| \frac{1}{\delta_1 - 1} - \frac{1}{\delta_2 - 1} \right|.$$

□

4 Monotonicity and Continuity Results.

Now we want to show some monotonicity results of the functions $i \to V_k(i)$. (It is then easy to derive monotonicity for $i \to c_k(i)$ resp. $i \to c_k'(i)$.) To prove monotonicity with respect to parameters of the prior distribution we need an appropriate concept of stochastic ordering. As it is easy to see that the function $s \to V_n(i,s)$ is increasing, a natural candidate is the usual stochastic order $\leq_{st}$. But this order relation is not preserved under conditioning. If one looks for a stronger order relation, which is preserved under conditioning, then one is lead in a natural way to the *likelihood ratio ordering*, cf. Keilson and Sumita (1982). Using this ordering resp. its multivariate counterpart $\leq_{tp}$, introduced by Karlin and Rinott (1980), it is possible to show general results about monotonicity for partially observed control models. For a detailed exposition of this approach see Rieder (1991). By applying Theorem 4.3 in that article to our model we get the following result.

Theorem 4.1 *Assume that*

(i) $(\theta, z) \to q(\theta, z)$ *is a* TP_2*-function.*

(ii) There are some $i, j \in I$*, such that* $\mu(i,\cdot) \leq_{tp} \mu(j,\cdot)$ *and* $\mu(i,\cdot)$ *or* $\mu(j,\cdot)$ *is* TP_2*.*

Then $c_k(i) \leq c_k(j)$ *resp.* $c_k'(i) \leq c_k'(j)$*.*

However, the assumption that $q(\theta, z)$ is a TP_2-function is often not fulfilled. It is well known that if q is twice differentiable, then q is TP_2 iff

$$\frac{d}{d\theta}\frac{d}{dz}\log q(\theta,z) \;\geq\; 0 \qquad \text{for all} \quad \theta, z.$$

But if e.g. $Q(\theta,\cdot) = \,,(\theta,\nu)$, then

$$\frac{d}{d\theta}\frac{d}{dz}\log q(\theta,z) \;=\; -1 \qquad \text{for all} \quad \theta, z,$$

and hence q is not TP_2.

Therefore, it is often better to prove monotonicity results directly by induction, using the fact that $s \to V_n(i,s)$ is increasing. We will do this now for the case of offers from a gamma distribution.

Theorem 4.2 *In Example 2.2 the function* $V_n(x,y,s)$ *is increasing in* x *and decreasing in* y *(in the model without recall as well as in the model with recall).*

Proof. By Example 3.1 c) $\mathrm{GB2}(\lambda, y, x, \cdot)$ is $\leq_{st}$-increasing in x and $\leq_{st}$-decreasing in y. Combining this with the fact that $s \to V_n(i,s)$ is increasing (in both models), the result follows from (2.5) and Theorem 3.2 a) by induction on n. □

Remark. Due to his computational results, Tamaki (1983) conjectured for the model without recall, that $y \to V_n(x,y,s)$ is decreasing. But he was not able to prove this theoretically except for $\lambda = 1$.

For the case of normally distributed offers we have the following result. We skip the proof, as it is similar to that of Theorem 4.2.

Theorem 4.3 *In Example 2.1 the function $V_n(x, b, s)$ is increasing in x (in the model without recall as well as in the model with recall).*

There is no general monotonicity result for $b \to V_n(x, b, s)$. In the model without recall $b \to V_n(x, b, s)$ is decreasing, if $\beta = 1$. This can be proved by using the representation for V_n given in Lemma 4.7. We conjecture that this is also true in case $\beta < 1$, but it seems to be difficult to prove this. In the model with recall, however, it is easy to see, that $b \to V_1(x, b, s)$ is strictly increasing if $s < -c + \beta x$.

Next we want to show that in Example 2.1 and 2.2 we have continuity of the function $i \to V_n(i, s)$. In fact we will show that this functions are locally Lipschitz, and we will give exact bounds for their oscillation. All subsequent results hold for the model with recall as well as for the model without recall. In the proofs we will restrict ourselves to the model with recall, as this model is the computationally more difficult one.

The main tool of our investigations will be the following lemma.

Lemma 4.4 *a) $\|V_n(i, \cdot)\|_L \le 1$ for all $i \in I$ and $n \in \mathbb{N}_0$.*
b) For every $f \in \mathfrak{L}$ and $P, Q \in \mathbb{P}_m$ it holds

$$\left| \int f \; dP - \int f \; dQ \right| \;\le\; \|f\|_L \cdot \zeta(P, Q).$$

Proof. a) follows from (2.1) by induction, using the fact that for any probability measure P and all Lipschitz functions it holds $\|\int P(dz) f(z, \cdot)\|_L \;\le\; \max_z \|f(z, \cdot)\|_L$ and $\|\max\{f, g\}\|_L \;\le\; \max\{\|f\|_L, \|g\|_L\}$.

b) If $f \in \mathfrak{L}$, then $f/\|f\|_L \in \mathfrak{L}_1$. Hence

$$\left| \int f \; dP - \int f \; dQ \right| \;=\; \|f\|_L \cdot \left| \int \frac{f}{\|f\|_L} \; dP - \int \frac{f}{\|f\|_L} \; dQ \right| \;\le\; \|f\|_L \cdot \zeta(P, Q)$$

$\square$

A. Normally distributed offers.

In Theorem 4.3 we have shown, that $x \to V_n(x, b, s)$ is increasing. Now we are able to show bounds for the growth of this function. We will use the abbreviation

$$\sigma_n(t) \;:=\; \sum_{\nu=0}^{n-1} t^\nu, \;\; t > 0.$$

Theorem 4.5 *For all $b \in \mathbb{R}_{>0}$, $s \in \mathbb{R}$ and $n \in \mathbb{N}_0$ we have $\|V_n(\cdot, b, s)\|_L \le \beta \sigma_n(\beta)$.*

Proof. We proceed by induction. The assertion obviously holds for $n = 0$. Hence assume that $\|V_n(\cdot, b, s)\|_L \le \beta\sigma_n(\beta) =: d_n$ for all $b \in \mathbb{R}_{>0}$ and $s \in \mathbb{R}$.

(i) Fix $b \in \mathbb{R}_{>0}$ and $s \in \mathbb{R}$ and define $g_1(x,z) := (x\nu + zb)/(\nu + b)$ and $g_2(z) := s \vee z$. Then we have $\|g_1(x,\cdot)\|_L = b/(\nu+b)$ and $\|g_2\|_L = 1$. Hence we get for

$$h(x,z) \ := \ V_n(g_1(x,z), \frac{\nu b}{\nu + b}, g_2(z)) \ = \ V_n\left(\frac{x\nu + zb}{\nu+b}, \frac{\nu b}{\nu+b}, s \vee z\right),$$

using the induction hypothesis and Lemma 4.4 a),

$$\|h(x,\cdot)\|_L \ \le \ d_n \cdot \|g_1\|_L + 1 \cdot \|g_2\|_L \ = \ d_n \cdot \frac{b}{\nu+b} + 1 \tag{4.1}$$

and $\|h(\cdot,z)\|_L \le d_n \cdot \nu/(\nu+b)$ for all $z \in \mathbb{R}$.

(ii) By Example 3.2 a) we have $\zeta(\mathcal{N}(x,b), \mathcal{N}(y,b)) = |x-y|$ for all $x, y \in \mathbb{R}$. Hence for $f(x) := \int \mathcal{N}(x, b+\nu, dz)\ h(x,z)$ we can conclude

$$\begin{aligned}
|f(x) - f(y)| \ &= \ \left| \int \mathcal{N}(x, b+\nu, dz)\ h(x,z) - \int \mathcal{N}(y, b+\nu, dz)\ h(y,z) \right| \\
&\le \ \int \mathcal{N}(x, b+\nu, dz)\ |h(x,z) - h(y,z)| \\
&\quad + \ \left| \int \mathcal{N}(x, b+\nu, dz)\ h(y,z) - \int \mathcal{N}(y, b+\nu, dz)\ h(y,z) \right| \\
&\le \ d_n \cdot \frac{\nu}{\nu+b} \cdot |x-y| \ + \ \|h(y,\cdot)\|_L \cdot \zeta(\mathcal{N}(x,b), \mathcal{N}(y,b)) \\
&= \ d_n \cdot \frac{\nu}{\nu+b} \cdot |x-y| \ + \ (d_n \cdot \frac{b}{\nu+b} + 1) \cdot |x-y| \\
&= \ (d_n + 1) \cdot |x-y|.
\end{aligned}$$

Applying this to $V_{n+1}(x,b,s) = \max\{s, -c + \beta f(x)\}$ yields

$$\|V_{n+1}(\cdot, b, s)\|_L \ \le \ \beta \cdot (d_n + 1) \ = \ \beta \cdot (\beta\sigma_n(\beta) + 1) \ = \ \beta \cdot \sigma_{n+1}(\beta).$$

□

For the sensitivity with respect to b we can show the following result.

Theorem 4.6 *For all $x, s \in \mathbb{R}$ and $b, b' \in \mathbb{R}_{>0}$ we have*

$$|V_n(x,b,s) - V_n(x,b',s)| \ \le \ \gamma_n \cdot |\sqrt{b} - \sqrt{b'}| + \rho_n \cdot |b - b'|, \tag{4.2}$$

where the sequences $(\gamma_n), (\rho_n)$ are defined recursively as follows:
$\gamma_0 = \rho_0 = 0$ and for $n \in \mathbb{N}_0$

$$\gamma_{n+1} \ = \ \beta \cdot \left(\gamma_n + \sigma_{n+1}(\beta) \cdot \sqrt{\frac{2}{\pi}}\right)$$

and

$$\rho_{n+1} \ = \ \beta \cdot \left(\rho_n + \beta\sigma_n(\beta) \cdot \sqrt{\frac{2}{\pi\nu}}\right).$$

Proof. As V_0 is independent of b, there is nothing to show for $n = 0$. Hence we assume that (4.2) holds for some $n \in \mathbb{N}_0$. For fixed $s, x \in \mathbb{R}$ we define

$$f_b(z) \;:=\; V_n\Big(\frac{x\nu + zb}{\nu + b}, \frac{\nu b}{\nu + b}, s \vee z\Big)$$

and $Q_b(\cdot) := \mathcal{N}(x, b + \nu)$.

(i) Combining Theorem 4.5 with the induction hypothesis yields

$$\begin{aligned}
|f_b(z) - f_{b'}(z)| \;&=\; \left|V_n\Big(\frac{x\nu + zb}{\nu + b}, \frac{\nu b}{\nu + b}, s \vee z\Big) - V_n\Big(\frac{x\nu + zb'}{\nu + b'}, \frac{\nu b'}{\nu + b'}, s \vee z\Big)\right| \\
&\le \left|V_n\Big(\frac{x\nu + zb}{\nu + b}, \frac{\nu b}{\nu + b}, s \vee z\Big) - V_n\Big(\frac{x\nu + zb'}{\nu + b'}, \frac{\nu b}{\nu + b}, s \vee z\Big)\right| \\
&\quad + \left|V_n\Big(\frac{x\nu + zb'}{\nu + b'}, \frac{\nu b}{\nu + b}, s \vee z\Big) - V_n\Big(\frac{x\nu + zb'}{\nu + b'}, \frac{\nu b'}{\nu + b'}, s \vee z\Big)\right| \\
&\overset{4.5}{\le} \beta\sigma_n(\beta) \cdot \left|\frac{x\nu + zb'}{\nu + b'} - \frac{x\nu + zb}{\nu + b}\right| \\
&\quad + \gamma_n \cdot \left|\sqrt{\frac{\nu b'}{\nu + b'}} - \sqrt{\frac{\nu b}{\nu + b}}\right| \;+\; \rho_n \cdot \left|\frac{\nu b'}{\nu + b'} - \frac{\nu b}{\nu + b}\right| \\
&\le \gamma_n \cdot \left|\sqrt{b'} - \sqrt{b}\right| \;+\; \rho_n \cdot |b' - b| \;+\; \beta\sigma_n(\beta) \cdot \frac{|x - z| \cdot |b - b'|}{\nu + b}.
\end{aligned}$$

(ii) An easy calculation shows $\int Q_b(dz)\,|x - z| = \sqrt{b + \nu} \cdot \sqrt{2/\pi}$, and by Example 3.2 b) we have $\zeta(Q_b, Q_{b'}) = \sqrt{2/\pi} \cdot |\sqrt{b + \nu} - \sqrt{b' + \nu}|$. Furthermore, from equation (4.1) we get $\|f_b\|_L \le \sigma_{n+1}(\beta)$. Hence we obtain

$$\begin{aligned}
|V_{n+1}(x, b, s) - V_{n+1}(x, b', s)| \;&\le\; \beta \cdot \left|\int Q_b(dz) f_b(z) - \int Q_{b'}(dz) f_{b'}(z)\right| \\
&\le \beta \cdot \left|\int Q_b(dz)|f_b(z) - f_{b'}(z)| \;+\; \Big(\int Q_b(dz) f_{b'}(z) - \int Q_{b'}(dz) f_{b'}(z)\Big)\right| \\
&\overset{(i),\,4.4}{\le} \beta \cdot \Big(\gamma_n \cdot \left|\sqrt{b'} - \sqrt{b}\right| \;+\; \rho_n \cdot |b' - b| \\
&\quad + \beta\sigma_n(\beta) \cdot \frac{|b - b'|}{\nu + b} \cdot \sqrt{b + \nu} \cdot \sqrt{2/\pi} \;+\; \zeta(Q_b, Q_{b'}) \cdot \|f_{b'}\|_L\Big) \\
&\le \beta \cdot \Big(\gamma_n \cdot \left|\sqrt{b'} - \sqrt{b}\right| \;+\; \rho_n \cdot |b' - b| \\
&\quad + \beta\sigma_n(\beta) \cdot |b - b'| \cdot \sqrt{\frac{2}{\pi\nu}} + \sqrt{\frac{2}{\pi}} \cdot |\sqrt{b} - \sqrt{b'}| \cdot \sigma_{n+1}(\beta)\Big) \\
&= \gamma_{n+1} \cdot |\sqrt{b} - \sqrt{b'}| \;+\; \rho_{n+1} \cdot |b - b'|.
\end{aligned}$$

$\square$

In case $\beta = 1$, the results of Theorem 4.5 and 4.6 can be improved considerably. This is due to the following lemma, which can easily be proved by induction.

Lemma 4.7 *If $\beta = 1$, then we have for all $x, s, t \in \mathbb{R},\ b \in \mathbb{R}_{\geq 0}$ and $n \in \mathbb{N}$:*

$$V_n(x, b, s) \ = \ t + V_n(x - t, b, s - t). \tag{4.3}$$

Theorem 4.8 *If $\beta = 1$, then $\|V_n(\cdot, b, s)\|_L \leq 1$ for all $s \in \mathbb{R},\ b \in \mathbb{R}_{\geq 0}$ and $n \in \mathbb{N}_0$.*

Proof. By Lemma 4.7 we have $V_n(x, b, s) = x + V_n(0, b, s - x)$. Combining this with Theorem 4.3 and the fact that $s \to V_n(0, b, s)$ is increasing yields the assertion. $\square$

Utilizing Theorem 4.8 we can also improve the result of Theorem 4.6. For a detailed proof see Müller (1995).

Theorem 4.9 *In case $\beta = 1$ for any $x, s \in \mathbb{R}$ and $b, b' \in \mathbb{R}_{>0}$ it holds*

$$|V_n(x, b, s) - V_n(x, b', s)| \ \leq \ \gamma_n \cdot |\sqrt{b} - \sqrt{b'}| + \rho_n \cdot |b - b'| \tag{4.4}$$

with $\gamma_n = 1/\sqrt{2\pi} \cdot n$ and $\rho_n = \sqrt{2/(\pi\nu)} \cdot n$.

A. Gamma distributed offers.

For gamma distributed offers we are able to show similar results. The proofs are skipped, since they are similar to the case of normal distributions. We refer to Müller (1995) for more details.

Theorem 4.10 *For all $y > 1,\ s, x, x' \in \mathbb{R}_{>0}$ and $n \in \mathbb{N}_0$ we have*

$$|V_n(x, y, s) - V_n(x', y, s)| \ \leq \ \frac{\beta\lambda}{y - 1} \cdot \sigma_n\left(\beta + \frac{\beta\lambda}{y - 1}\right) \cdot |x - x'|. \tag{4.5}$$

Theorem 4.11 *If $y, y' \geq \alpha > 1$, then we have for all $x, s \in \mathbb{R}_{>0}$ and $n \in \mathbb{N}_0$:*

$$|V_n(x, y, s) - V_n(x, y', s)| \ \leq \ \gamma_n \cdot x \cdot \left|\frac{1}{y - 1} - \frac{1}{y' - 1}\right|, \tag{4.6}$$

where the sequence (γ_n) is recursively defined as follows: $\gamma_0 = 0$ and

$$\gamma_{n+1} \ = \ \beta \cdot \left(\gamma_n + \lambda \cdot \left(\frac{\gamma_n}{\alpha - 1} + \frac{\beta\lambda}{\alpha - 1} \cdot \sigma_n\Big(\beta + \frac{\beta\lambda}{\alpha - 1}\Big) + 1\right)\right).$$

5 Numerical Results.

In the model without recall, the value function V_k and the control limit function c_k can be computed numerically. This is due to the fact, that $s \to V_k(i, s)$ if of a very simple structure, namely the maximum of two affine functions. The explicit formulas can be found in DeGroot (1968) and Tamaki (1983). The latter article also includes some lists with numerical values.

In the model with recall, however, this is not the case. Therefore the model with recall is numerically intractable, and hence it is of special interest to investigate the structure of the solution as we did in the previous section.

For the (computationally tractable) model without recall and with normally distributed offers we computed in case $\beta = 1$ the exact values of the function $c_k(0, b)$ and compared the numerical results with the analytical bounds for the oscillation given in Theorem 4.9. It turned out that - especially for large N - the analytical bounds are not very tight, but this is not surprising due to the recursive nature of the bounds.

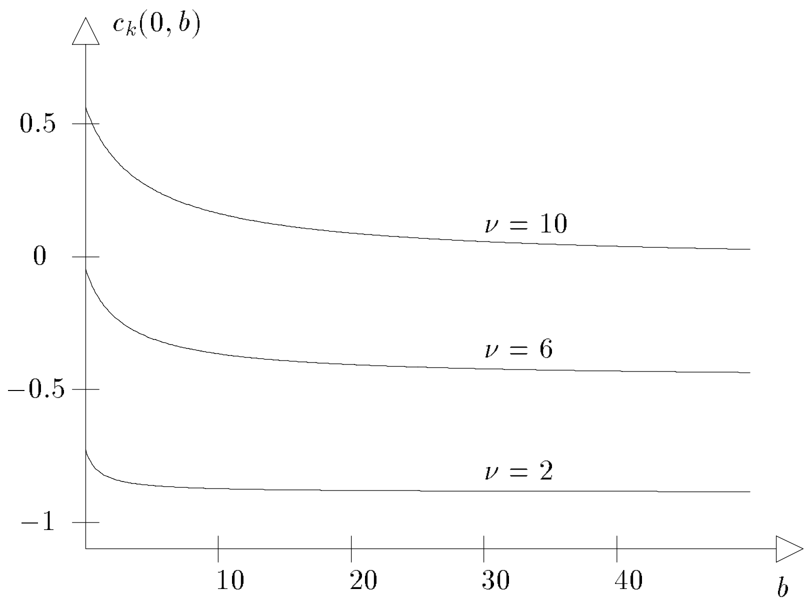

Figure 1: $b \to c_k(0, b)$ for $\beta = 1,\ c = 1,\ k = 40$ and $\nu = 2, 6$ and 10.

From Figure 1 you can see, however, that the qualitative behavior of $b \to c_k(0,b)$ is described very well by (4.4). In fact, $b \to c_k(0,b)$ seems not to be Lipschitz in a neighbourhood of 0. On the contrary, it somehow looks like $b \to \alpha_1 - \alpha_2 \cdot \sqrt{b}$ for some constants $\alpha_1 \in \mathbb{R}$ and $\alpha_2 > 0$, as one would expect if the bounds in (4.4) where tight. By rescaling the b-axe with the transformation $b \to \sqrt{b}$ one gets a function, which is Lipschitz in a neighbourhood of 0, as claimed in Theorem 4.9, cf. Figure 2.

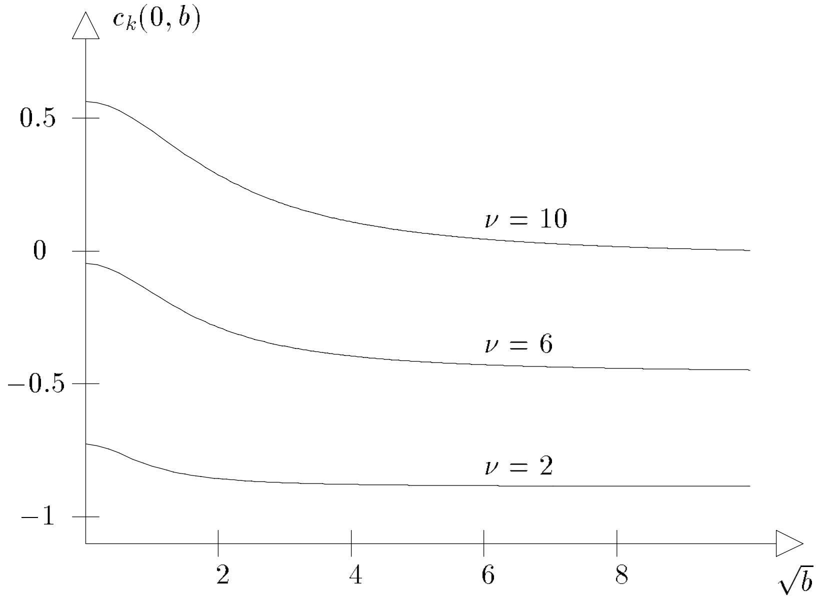

Figure 2: $\sqrt{b} \to c_k(0,b)$ for $\beta = 1,\ c = 1,\ k = 40$ and $\nu = 2, 6$ and 10.

Acknowledgements

The content of this paper appears in part in my dissertation Müller (1995). I wish to express my sincere gratitude to K. Hinderer for his expert guidance and encouragement in the course of the development of that work.

References

CUMMINS J.D., G. DIONNE, J.B. MCDONALD, B.M. PRITCHETT (1990). Applications of the GB2 Family of Distributions in Modeling Insurance Loss Processes. *Insurance: Mathematics and Economics* **9**, 257 - 272.

DEGROOT M. H. (1968). Some Problems of optimal stopping. *J. Roy. Stat. Soc. Series B* **30**, 108 - 122.

HINDERER K. (1970). *Foundations of Non-stationary Dynamic Programming with Discrete Time Parameter.* Lecture Notes in Oper. Res. and Math. Syst. **33**. Springer, Berlin.

KARLIN S., Y. RINOTT (1980). Classes of Orderings of Measures and Related Correlation Inequalities. *Journal of Multivariate Analysis* **10**, 467 - 498.

KEILSON J., U. SUMITA (1982). Uniform stochastic ordering and related inequalities. *Canad. J. Statist.* **10**, 181 -198.

LEHMANN E. (1955). Ordered families of distributions. *Ann. Math. Statist.* **26**, 399 - 419.

MÜLLER A. (1995). *Ordnungen und Metriken auf Mengen von Wahrscheinlichkeitsmaßen mit Anwendungen bei Markoffschen Entscheidungsprozessen.* Dissertation, Universität Karlsruhe.

RACHEV S.T. (1991). *Probability Metrics and the Stability of Stochastic Models.* Wiley.

RIEDER U. (1975). Bayesian Dynamic Programming. *Adv. Appl. Prob.* **7**, 330 - 348.

RIEDER U. (1988). *Bayes'sche Kontrollmodelle.* Skriptum, Universität Ulm.

RIEDER U. (1991). Structural Results for partially observed control models. *Zeitschrift f. Oper. Res.* **35**, 473 - 490.

RIEDER U., H. WAGNER (1991). Structured policies in the sequential design of experiments. *Annals of Oper. Res.* **32**, 165 - 188.

SAKAGUCHI M. (1961). Dynamic Programming of some sequential sampling design. *J. Math. Anal. Appl.* **2**, 446 - 466.

SHAKED, M., J.G. SHANTHIKUMAR (1994). *Stochastic Orders and their Applications.* Academic Press, London.

STEWART T.J. (1978). Optimal Selection from a random sample with learning of the underlying distribution. *J. Amer. Stat. Asoc.* **73**, 775 - 780.

TAMAKI M. (1983). Optimal Selection from a Gamma Distribution with unknown Parameter. *Zeitschrift f. Oper. Res.* **28**, 47 - 57.

ZOLOTAREV V.M. (1983). Probability metrics. *Theory Prob. Appl.* **28**, 278-302.