ECON5160: Problems solved

This version current as of April 1st (for those with Math 4 credits: see separate note for the EVT problems); a correction April 2nd to page 21 (plus minor tweaks).

Problems for February 06

Schweder #9 If each («father», «son») = (X, Y) pair is bivariate normal (Gaussian), then the conditional density $f_{Y|X}(y, x)$ is given as the joint distribution divided by the marginal. With common marginals, X and Y have common mean μ and common stdev σ, so we get

$$f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)} = \frac{(2\pi\sigma^2\sqrt{1-\rho^2})^{-1}\exp\Big(-\frac{(x-\mu)^2-2\rho(x-\mu)(y-\mu)+(y-\mu)^2}{2\sigma^2(1-\rho^2)}\Big)}{(\sigma\sqrt{2\pi})^{-1}\exp\Big(-\frac{(x-\mu)^2}{2\sigma^2}\Big)}$$

which after brief calculations turn into

$$= (\sigma\sqrt{2\pi(1-\rho^2)})^{-1}\cdot\exp\Big(-\frac{(y-\mu(1-\rho)-\rho x)^2}{2\sigma^2(1-\rho^2)}\Big).$$

So we assume our Markov chain X_n to have a transition density kernel $K(x, y) = f_{X_n|X_{n-1}}(y|x)$ defined by the latter expression – a Gaussian. This solves the «construct» part of the problem.
What is a stationary distribution? If a probability density $g(x)$ satisfies

$$\int_{\mathbb{R}} K(x,y)g(x)\,dx = g(y),$$

then the sons will have the same (unconditional) distribution, and by the Markov property their sons the same etc. – precisely the definition of a stationary distribution. We are asked to show that the marginal distribution $f_X(x)$ satisfies this. However, then the left hand side is by construction

$$\int_{\mathbb{R}} \frac{f_{X,Y}(x,y)}{f_X(x)} f_X(x)\,dx = \int_{\mathbb{R}} f_{X,Y}(x,y)\,dx = f_Y(y),$$

the marginal distribution of Y, which by assumption is the same as $f_X(x)$, which solves the «show» part.

TK III P. 1.1 We have N persons, and X_n is the number infected by time n. At step n, two are drawn at random from the $N \cdot (N-1)$ possible pairs; we have {both equal}={both infected} or {none infected} (disjoint union), so

$$\begin{aligned} q = q_x = \mathsf{Pr}[\text{same status}|X_n = x] &= \frac{x}{n} \cdot \frac{x-1}{N-1} + \frac{N-x}{N} \cdot \frac{N-x-1}{N-1} \\ &= \frac{x(x-1) + (N-x)(N-x-1)}{N(N-1)} \end{aligned}$$

in which case $X_{n+1} = X_n$; this also occurs in $1-\alpha$ of the remaining cases. So in the language of transition probabilities, we have

$$\begin{aligned} p_{xx} &= q_x + (1-\alpha)(1-q_x) = 1 - \alpha(1-q_x) \\ p_{x,x+1} &= 1 - p_{xx} = \alpha(1-q_x). \end{aligned}$$

TK III P. 1.4 Answer:

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \end{array} \begin{array}{c} \begin{array}{cccc} 0 & 1 & 2 & 3 \end{array} \\ \begin{bmatrix} \frac{1}{10} & \frac{3}{10} & \frac{2}{10} & \frac{4}{10} \\ 0 & \frac{4}{10} & \frac{2}{10} & \frac{4}{10} \\ 0 & 0 & \frac{6}{10} & \frac{4}{10} \\ 0 & 0 & 0 & 1 \end{bmatrix} \end{array}$$

TK III P. 2.1 Observe that $\mathbf{P}$ is doubly stochastic – and as such, as we know by now (ch. IV) has a uniform stationary distribution.

TK III P. 2.4 (This is a «process with memory» augmented with yesterday's state to obtain the Markov property.)

$$\mathbf{P} = \begin{array}{c} \\ 00 \\ 01 \\ 10 \\ 11 \end{array} \begin{array}{c} \begin{array}{cccc} 00 & 01 & 10 & 11 \end{array} \\ \begin{bmatrix} \alpha & 1-\alpha & 0 & 0 \\ 0 & 0 & 1-\beta & \beta \\ \alpha & 1-\alpha & 0 & 0 \\ 0 & 0 & 1-\beta & \beta \end{bmatrix} \end{array}$$

TK III P. 2.5 If $\mathbf{P}^3 = \begin{bmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ 0 & 0 & 1 \end{bmatrix}$ then $\mathsf{Pr}[X_3 = 0|X_0 = i, T > 3] = \frac{q_{i0}}{q_{i0}+q_{i1}}$, for $i = 0, 1$.

TK III P. 3.10 (This problem was actually stated by mistake. The main «difficulty» here is the ambiguity in the language as to whether the «instantly produced» quantity is added «at the end of period $n-1$» or «just before period n». Since X_n is interpreted as «inventory

quantity on hand», I will not allow for $X_n < 0$, but – maybe a bit inconsistent – I choose to allow for $X_n < s$ in order to be able to read any deficit under s off the value of X itself. In part (b) we are asked for p_{04}, so evidently we are supposed to allow for $X = 0$. So the interpretation is that if $X_n \leq s$, then $X_{n+1} = \max\{0, S - \xi_{n+1}\}$; if $X_n > s$ then $X_{n+1} = \max\{0, X_n - \xi_{n+1}\}$.)
Part (a) can be calculated by hand. For part (b), we observe that if $X_n = 0$ then by restocking we will immediately have 4, so $X_{n+1} = 4$ if and only if $\xi_{n+1} = 0$, which happens with probability 0.1. An p_{41}, the transition from $X_n = 4$ to $X_{n+1} = 1$, occurs if and only if $\xi_{n+1} = 3$, which happens with probability 0.2.

TK III P. 4.7 The law of double expectation is key:

$$\begin{aligned}
h(X_0) &= \mathsf{E}[\sum_{n=0}^{\infty} \beta^n c(X_n)|X_0] \\
&= \beta^0 c(X_0) + \mathsf{E}\big[\mathsf{E}[\sum_{n=1}^{\infty} \beta^n c(X_n)|X_1]\big|X_0\big] \\
&= c(X_0) + \beta\mathsf{E}\big[\underbrace{\mathsf{E}[\sum_{m=0}^{\infty} \beta^m c(X_{m+1})|X_1]}_{=h(X_1) \text{ by the Markov property}}\big|X_0\big] \\
&= c(X_0) + \beta\mathsf{E}[h(X_1)|X_0] \\
&= c(X_0) + \beta\sum_j P_{X_0,j} h(j),
\end{aligned}$$

valid for each possible initial state $X_0 = i$.

TK III P. 4.17 One cannot reach $X = 2$ in a single step, so define $\psi_i(s) = \mathsf{E}[s^T|X_0 = i]$. Using general probabilities except for the zeroes inserted explicitly,

$$\begin{aligned}
\psi_1(s) &= P_{12}s^1 + P_{11}s\psi_1(s) + 0, \qquad \text{so that} \\
\psi_1(s) &= \frac{sP_{12}}{1 - sP_{11}}
\end{aligned}$$

(the assumption $s < 1$ guarantees nonzero denominator). For $i = 0$,

$$\begin{aligned}
\phi(s) = \psi_0(s) &= \mathsf{E}[s^T|X_0 = 0] = \mathsf{E}[\mathsf{E}[s^T|X_0 = 0, X_1]|X_0 = 0] \\
&= s\psi_0(s) + s\psi_1(s), \qquad \text{so that} \\
\phi(s) &= \frac{s}{1-s}\psi_1(s) = \frac{s^2 P_{12}}{(1-s)(1-sP_{11})}.
\end{aligned}$$

Now insert for the probabilities.

TK III P. 7.1 The transition matrix is

$$\mathbf{P} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \\ m \\ \vdots \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & 3 & \cdots & m-1 & \cdots \end{array} \\ \left[\begin{array}{ccccccc} 1 & 0 & 0 & 0 & \cdots & 0 & \cdots \\ 1 & 0 & 0 & 0 & \cdots & 0 & \cdots \\ 1/2 & 1/2 & 0 & 0 & \cdots & 0 & \cdots \\ 1/3 & 1/3 & 1/3 & 0 & \cdots & 0 & \cdots \\ 1/4 & 1/4 & 1/4 & 1/4 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 & \cdots \\ 1/m & 1/m & 1/m & 1/m & \cdots & 1/m & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] \end{array}$$

Now the «**I**» in the decomposition is the top-left «1», so **Q** is formed by removing the top row and leftmost column (both «#0»):

$$\mathbf{I} - \mathbf{Q} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ \vdots \\ m \\ \vdots \end{array} \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 3 & \cdots & m-1 & \cdots \end{array} \\ \left[\begin{array}{cccccc} 1 & 0 & 0 & \cdots & 0 & \cdots \\ -1/2 & 1 & 0 & \cdots & 0 & \cdots \\ -1/3 & -1/3 & 1 & \cdots & 0 & \cdots \\ -1/4 & -1/4 & -1/4 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & 1 & \cdots \\ -1/m & -1/m & -1/m & \cdots & -1/m & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right] \end{array}$$

It certainly helps to notice that if this were the coefficient matrix of a linear equation system, we could solve it from above; the inverse will also be lower triangular, and if one puts up, say, the 4×4 case one should see a pattern which could then be proved by induction:

$$(\mathbf{I} - \mathbf{Q})^{-1} = \left[\begin{array}{cccccc} 1 & 0 & 0 & \cdots & 0 & \cdots \\ 1/2 & 1 & 0 & \cdots & 0 & \cdots \\ 1/2 & 1/3 & 1 & \cdots & 0 & \cdots \\ 1/2 & 1/3 & 1/4 & \cdots & 0 & \cdots \\ \vdots & \vdots & \vdots & \ddots & 1 & \cdots \\ 1/2 & 1/3 & 1/4 & \cdots & 1/(m-1) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{array}\right]$$

(Conditional on X_0 we only need the X_0-dimensional version, but if X_0 is drawn from a distribution on the nonnegative integers, we might as well take the infinite-dimensional version of the expression.)

Problems for February 13

TK IV P. 1.3 There are (at least) two methods available, one by solving the eigenvalue problem and one by expected first return-time from TK IV.4; we will cover both.
The eigenvalue problem approach: Observe first that the α_i are indexed from $i = 1$ to 6 although the states are indexed from 0 to 5. If $\boldsymbol{\pi}$ is the limiting distribution, it solves the eigenvalue problem

$$\mathbf{0} = \boldsymbol{\pi}(\mathbf{P} - \mathbf{I})$$

where the right hand side is arguably clearer in colum form, as the transpose of

$$\begin{bmatrix} \pi_0\alpha_1 - \pi_0 + \pi_1 \\ \pi_0\alpha_2 - \pi_1 + \pi_2 \\ \pi_0\alpha_3 - \pi_2 + \pi_3 \\ \pi_0\alpha_4 - \pi_3 + \pi_4 \\ \pi_0\alpha_5 - \pi_4 + \pi_5 \\ \pi_0\alpha_6 - \pi_5 \end{bmatrix}$$

Now solve this for $\pi_1, \cdots, \pi_5$ in terms of π_0:

$$\begin{aligned} \pi_1 &= \pi_0(1 - \alpha_1) \\ &= \pi_0(\alpha_2 + \cdots + \alpha_6) \\ \pi_2 &= \pi_1 - \pi_0\alpha_2 \\ &= \pi_0(\alpha_3 + \cdots + \alpha_6), \qquad \text{and similarly} \\ \pi_3 &= \pi_0(\alpha_4 + \cdots + \alpha_6) \\ \pi_4 &= \pi_0(\alpha_5 + \alpha_6) \\ \pi_5 &= \pi_0 \cdot \alpha_6 \end{aligned}$$

Now $\pi_1 + \cdots + \pi_5 = 1 - \pi_0$. We add up and solve for π_0 to get

$$\pi_0 = (\alpha_1 + 2\alpha_2 + \cdots + 5\alpha_5)^{-1}$$

The basic limit theorem approach: In the language of TK p. 246, we learn from Theorem 4.1 and the remark on page 254 (by the existence of a limiting distribution given in the problem) we have

$$\begin{aligned} \pi_0^{-1} = m_0 &= \mathsf{E}[\text{first return time to } 0 | X_0 = 0] \\ &= \sum_{n=1}^{\infty} n\mathsf{Pr}[\text{first return time to 0 equals } n | X_0 = 0] \\ &= 1 \cdot \alpha_1 + 2 \cdot \alpha_2 \cdot 1 + \cdots + 5\alpha_5 \cdot 1^4 \end{aligned}$$

by inspecting the possible paths of X starting from 0.

TK IV P. 1.10 The matrix is doubly stochastic, so the limiting distribution is uniform: $\pi_i = 1/(N+1)$.

TK IV P. 1.12

(a) We denote by $\mathbf{\Pi}$ the square matrix whose rows are a stationary distribution $\boldsymbol{\pi}$, and we are supposed to show

$$\mathbf{P}^n = \mathbf{\Pi} + (\mathbf{P} - \mathbf{\Pi})^n.$$

Observe first for the case $n = 2$ that

$$\mathbf{\Pi} + (\mathbf{P} - \mathbf{\Pi})^2 = \mathbf{\Pi} + \mathbf{P}^2 - \mathbf{P\Pi} - \mathbf{\Pi P} + \mathbf{\Pi}^2$$

so we must establish the identity

$$\mathbf{0} = \mathbf{\Pi} - \mathbf{P\Pi} - \mathbf{\Pi P} + \mathbf{\Pi}^2$$

Now $\mathbf{\Pi}$ is itself a transition matrix with $\boldsymbol{\pi}$ as stationary distribution, so $\mathbf{\Pi}^2 = \mathbf{\Pi}$ (i.e. $\mathbf{\Pi}$ is *idempotent*). Furthermore, by the (left) eigenvector property of $\boldsymbol{\pi}$, we have $\mathbf{\Pi P} = \mathbf{\Pi}$ – indeed, those two properties show the orthogonality $\mathbf{\Pi}(\mathbf{P} - \mathbf{\Pi}) = \mathbf{0}$. Also, it turns out that we have a right eigenvector property: each column of $\mathbf{P}i$ is a constant times $\mathbf{1}$, and $\mathbf{P1} = \mathbf{1}$ since each row of $\mathbf{P}$ sums up to 1. Therefore $\mathbf{P\Pi} = \mathbf{\Pi}$ too, which also shows the orthogonality $(\mathbf{P} - \mathbf{\Pi})\mathbf{\Pi} = \mathbf{0}$. Now assume for induction that for some k,

$$\mathbf{P}^k = \mathbf{\Pi} + (\mathbf{P} - \mathbf{\Pi})^k.$$

Then

$$\mathbf{P}^{k+1} = \mathbf{P}^k\mathbf{P} = \mathbf{\Pi P} + (\mathbf{P} - \mathbf{\Pi})^k\mathbf{P}$$

and the proof will be complete if we can establish $(\mathbf{P} - \mathbf{\Pi})^k\mathbf{P} = (\mathbf{P} - \mathbf{\Pi})^k(\mathbf{P} - \mathbf{\Pi})$. However, this is equivalent to proving that $(\mathbf{P} - \mathbf{\Pi})^k\mathbf{\Pi} = \mathbf{0}$, which is true by the latter orthogonality property established.

(b) By e.g. the eigenvector approach, we find the stationary distribution $\boldsymbol{\pi} = \frac{1}{4}[1, 2, 1]$. Then $(\mathbf{P} - \mathbf{\Pi}) = \frac{1}{4}\begin{bmatrix} 1 & 0 & -1 \\ 0 & 0 & 0 \\ -1 & 0 & 1 \end{bmatrix}$ whose nth power is $n4^{-n}(\mathbf{P} - \mathbf{\Pi})$, from which $\mathbf{P}^n$ easily follows.

TK IV P. 3.2 We have irreducibility: all states (i, j) commute, so that for each i and j, j is accessible from i in n_{ij} steps. We have aperiodicity: for each i, $(\mathbf{P}^n)_{ii} > 0$ for all n greater than some n_i. So for all $n > \nu_{ij} := n_{ij} + n_j$, we have $(\mathbf{P}^n)_{ij} > 0$. Since we have a finite number of states, $\bar{n} := \max_{i,j} \nu_{ij}$ is finite, and $\mathbf{P}^n$ has only positive entries whenever $n > \bar{n}$. This proves regularity.
Now the expected number of hits for a given transient state, is finite (see the proof of TK Theorem 3.1). Therefore, since we have a finite number of states, at least one is recurrent. By corollary 3.1, irreducibility implies that all states are recurrent. We are done.

TK IV P. 4.1

(a) $\left[\frac{\beta}{\alpha+\beta} \quad , \quad \frac{\alpha}{\alpha+\beta}\right]\mathbf{P} = \left[\frac{\beta-\alpha\beta+\alpha\beta}{\alpha+\beta} \quad , \quad \frac{\alpha\beta+\alpha-\alpha\beta}{\alpha+\beta}\right]$ and $\frac{\beta}{\alpha+\beta} + \frac{\alpha}{\alpha+\beta} = 1$.

(b) For $n > 2$, first return in n steps is the event «$011\cdots10$» with $n-2$ ones.

(c) We have $m_0 = 1 - \alpha + \alpha\sum_{n=2}^{\infty} n\beta(1-\beta)^{n-2}$.
This calls for summation by parts, which you may not be too familiar with: To fit it into the notation of `http://en.wikipedia.org/wiki/Summation_by_parts#Method`, rewrite the partial sums into

$$\sum_{n=0}^{N}(n+2)\beta(1-\beta)^n$$

and put $a_n = n + 2$ and $B_n = \sum_{k=0}^{n}\beta(1-\beta)^k = 1-(1-\beta)^{n+1}$. Then $a_{n+1} - a_n = 1$ so that the sum becomes

$$(N+2)(1-(1-\beta)^{N+1}) - \sum_{n=0}^{N-1}(1-(1-\beta)^{n+1})$$
$$= 2 - (N+2)(1-\beta)^{N+1} + \sum_{n=0}^{N-1}(1-\beta)^{n+1}$$

which converges to $2 + (1-\beta)/\beta = 1 + 1/\beta$ so that

$$m_0 = 1 + \alpha/\beta = \frac{\beta+\alpha}{\beta}$$

which we recognize as π_0^{-1}.

TK IV P. 4.4 This is basically the possibly infinite-dimensional version of problem 1.3, except we are not given the information that a limiting distribution exists. Solving like in problem 1.3, we get that the vector $\boldsymbol{\pi}$ with components $\pi_0 = (\sum_k k\alpha_k)^{-1}$, and $\pi_i = \pi_0\sum_{j>i}\alpha_j$ for $i > 0$, satisfies the eigenvector equation. It will be a probability distribution, and hence a stationary distribution, if and only if $\sum_k k\alpha_k$ converges.

To show that this is a limiting distribution, assume first that the state space is «truly infinite» in the sense that for any k there is an $\alpha_i > 0$ for some $i > k$. Then the process is recurrent, irreducible and aperiodic, and theorems 4.1 and 4.2 grants that $\boldsymbol{\pi}$ is a limiting distribution as long as if $\sum_k k\alpha_k$ converges (which ensures positive recurrence). If on the other hand there is a greatest k such that $\alpha_k > 0$ but $\alpha_i = 0$ for all $i > k$ (so that all states $k, k+1, \ldots$ are transient), then observe first that $\sum_k k\alpha_k$ does converge; solve the k-state problem for the limiting distribution, which then exists for states $0, \ldots, k-1$; observe that putting the components $\pi_i = 0$ for $i \geq k$ indeed makes $\boldsymbol{\pi}$ a limiting distribution also for the full matrix.

Problems for February 20

1

(a) Seierstad 1.11:

$$\max \mathsf{E}[\sum_{t=0}^{T-1} -u_t^2 - X_T^2] \qquad \text{where } X_{t+1} = X_t V_{t+1} + u_t, \qquad u_t \in (-\infty, \infty)$$

We are given a hint which may be re-written as $J(t, x, v) = -c_t(v)x^2$, which is true for $t = T$ with $c_T = 1$. We proceed to prove the induction step:

$$\begin{aligned} J(t-1, x, V_{t-1}) &= \max_u \{-u^2 - \mathsf{E}[c_t(V_t)(xV_t + u_t)|V_{t-1}]\} \\ &= \max_u \{-u^2 - \mathsf{E}[c_t(V_t)(x^2V_t^2 + 2uxV_t + u^2)|V_{t-1}]\} \end{aligned}$$

We enhance the induction hypothesis with the additional hypothesis $\mathsf{E}[c_t(V_t)|V_{t-1}]+1 > 0$ (which is true for $t = T$), ensuring that a unique maximum exists. The maximand is of the form $-\alpha u^2 - 2\beta u - \gamma$, with max for $u^* = -\beta/\alpha$ and maximum value $\beta^2/\alpha - \gamma$. This gives

$$J(t-1, x, V_{t-1}) = x^2 \frac{\mathsf{E}[c_t(V_t)V_t|V_{t-1}]^2}{1 + \mathsf{E}[c_t(V_t)|V_{t-1}]} - x^2 \mathsf{E}[c_t(V_t)V_t^2|V_{t-1}].$$

In our case, $V \in \{0, 1\}$ so that $V^2 = V$ and we get

$$\begin{aligned} J(t-1, x, V_{t-1}) &= x^2 \left(\frac{(c_t(1)\mathsf{Pr}[V_t = 1|V_{t-1}])^2}{1 + \mathsf{E}[c_t(V_t)|V_{t-1}]} - c_t(1)\mathsf{Pr}[V_t = 1|V_{t-1}]\right) \\ J(t-1, x, V_{t-1}) &= -x^2 \frac{c_t(1)\mathsf{Pr}[V_t = 1|V_{t-1}](1 + c_t(0)\mathsf{Pr}[V_t = 0|V_{t-1}])}{1 + c_t(0)\mathsf{Pr}[V_t = 0|V_{t-1}] + c_t(1)\mathsf{Pr}[V_t = 1|V_{t-1}]} \\ &= -x^2 c_{t-1}(V_{t-1}) \end{aligned}$$

which shows both the form and the additional positivity hypothesis. Now we can insert for the probabilities to get the actual values for $c_{t-1}(0)$ and $c_{t-1}(1)$ (which in Seierstad's notation are b_{t-1}, a_{t-1}, respectively) determined recursively:

$$\begin{aligned} c_{t-1}(0) &= \frac{c_t(1)\mathsf{Pr}[V_t = 1|V_{t-1} = 0](1 + c_t(0)\mathsf{Pr}[V_t = 0|V_{t-1} = 0])}{1 + c_t(0)\mathsf{Pr}[V_t = 0|V_{t-1} = 0] + c_t(1)\mathsf{Pr}[V_t = 1|V_{t-1} = 0]} \\ &= \frac{c_t(1)(1 + \frac{3}{4}c_t(0))}{4 + 3c_t(0) + c_t(1)} \qquad \text{and similarly,} \\ c_{t-1}(1) &= \frac{3c_t(1)(1 + \frac{1}{4}c_t(0))}{4 + c_t(0) + 3c_t(1)} \end{aligned}$$

(b) Seierstad 1.14:
We are asked to calculate the performance associated with $u_* = (1-\rho)$ and verify that it is equal to $(1-\rho)^{-\gamma}x_0^{1-\gamma}$, where $\rho^\gamma = \beta\mathsf{E}[V^{1-\gamma}]$ (is a constant). The performance is

$$\mathsf{E}\Big[\sum_{t=0}^{\infty}\beta^t x_t^{1-\gamma}u_{*t}^{1-\gamma}\Big] = (1-\rho)^{1-\gamma}\sum_{t=0}^{\infty}\beta^t\mathsf{E}[x_t^{1-\gamma}]$$

where $x_t = V_t x_{t-1}(1-u_{*,t-1}) = V_t V_{t-1}\cdot\dots\cdot V_1 x_0 \rho^t$, which by independence of the V's yields

$$\begin{aligned}\mathsf{E}\Big[\sum_{t=0}^{\infty}\beta^t x_t^{1-\gamma}u_{*t}^{1-\gamma}\Big] &= (1-\rho)^{1-\gamma}\sum_{t=0}^{\infty}\beta^t x_0^{1-\gamma}\mathsf{E}[V_t^{1-\gamma}]\cdot\dots\cdot\mathsf{E}[V_1]\cdot(\rho^{1-\gamma})^t \\ &= (1-\rho)^{1-\gamma}x_0^{1-\gamma}\sum_{t=0}^{\infty}(\beta\mathsf{E}[V^{1-\gamma}]\rho^{1-\gamma})^t.\end{aligned}$$

Now recall $\beta\mathsf{E}[V^{1-\gamma}]\rho^{-\gamma} = 1$, and sum up the geometric series $\sum\rho^t$ to get

$$\begin{aligned}&= (1-\rho)^{1-\gamma}x_0^{1-\gamma}(1-\rho)^{-1} \\ &= (1-\rho)^{-\gamma}x_0^{1-\gamma}\end{aligned}$$

which is what we were supposed to prove.

(c) Seierstad 1.15:
Note first that by $x_0 > 0$, we have all $X_t > 0$. The probability distribution of V is x-dependent, and we will turn out to need $\mathsf{E}_x[V]$, which equals x. For $t = T$ we have $u_T = 0$ and $J_T(x) = x$. We then have

$$\begin{aligned}J_{T-1}(x) &= \max_{u\in[0,1]}\{(1-u)x + \mathsf{E}_x[x + ux + V_T]\} \\ &= \max_{u\in[0,1]}\{2x + \mathsf{E}_x[V_T]\} \\ &= 3x \qquad \text{with any control } u_{T-1}\in[0,1] \text{ optimal}\end{aligned}$$

and furthermore,

$$\begin{aligned}J_{T-2}(x) &= \max_{u\in[0,1]}\{(1-u)x + 3\mathsf{E}_x[x + ux + V_{T-1}]\} \\ &= \max_{u\in[0,1]}\{4x + 2ux + 3\mathsf{E}_x[V_{T-1}]\} \\ &= 9x \qquad \text{with } u_{T-2}^* = 1.\end{aligned}$$

We guess for induction the form $A_t x$, and then

$$\begin{aligned}J_{t-1}(x) &= \max_{u\in[0,1]}\{(1-u)x + A_t\mathsf{E}_x[x + ux + V_t]\} \\ &= \max_{u\in[0,1]}\{(1+A_t)x + (A_t-1)u + A_t\mathsf{E}_x[V_t]\}.\end{aligned}$$

Now if $A_t \geq 1$, then $u^*_{t-1} = 1$ is optimal (and strictly optimal if strict inequality), in which case we get $3A_t x$, completing the induction proof: starting from $A_T = 1$, it follows that

$$\begin{aligned} J_{T-\tau}(x) &= 3^\tau x && \text{where } 3^\tau \geq 1 \text{ for } \tau \geq 0\text{, so} \\ u^*_{T-\tau} &= 1 && \text{is optimal for all } \tau > 0 \text{ (uniquely so for } \tau > 1\text{), while} \\ u^*_T &= 0. \end{aligned}$$

 # 2 Dynamics: $X_{t+1} = (X_t - c_t)(1 + r + \mathbf{u}_t \mathbf{V}_{t+1})$.

(a) Finite-horizon case. We are also asked how crucial it is to assume $\mathbf{V}_{t+1}|\mathbf{V}_t$ time-invariant; the following calculation will not assume it is, and I will comment on it at the end.

$$J_T(x, \mathbf{v}) = A_T \frac{x^{1-\gamma}}{1-\gamma} \qquad \text{with } c_T = X_T\text{, arbitrary portfolio}$$

$$\begin{aligned} J_{T-1}(x, \mathbf{v}) &= \sup_{c,\mathbf{u}} \{A_{T-1} \frac{c^{1-\gamma}}{1-\gamma} + \mathsf{E}_{T-1}\big[A_T \frac{\big((x-c)(1+r+\mathbf{u}\mathbf{V}_T)\big)^{1-\gamma}}{1-\gamma}\big]\} \\ &= \sup_c \{A_{T-1} \frac{c^{1-\gamma}}{1-\gamma} + (x-c)^{1-\gamma} \sup_{\mathbf{u}} \mathsf{E}_{T-1}\big[A_T \frac{(1+r+\mathbf{u}\mathbf{V}_T)^{1-\gamma}}{1-\gamma}\big]\} \end{aligned}$$

The optimization with respect to $\mathbf{u}$ does not depend on x (by itself, nor through c which might in optimum depend on x) – this will show two-fund monetary separation. Write the $\sup_{\mathbf{u}}$ part as $\eta_{T-1}/(1-\gamma)$ (then $\eta_{T-1} = \eta_{T-1}(\mathbf{v})$ is > 0)[1], to get

$$= \sup_c \{\frac{A_{T-1} c^{1-\gamma} + \eta_{T-1}(x-c)^{1-\gamma}}{1-\gamma}\}$$

This is a concave problem which may be solved explicitly; the solution is of the form $c^* = k^* x$, and we get

$$J_{T-1}(x, \mathbf{v}) = K_{T-1}(\mathbf{v}) \frac{x^{1-\gamma}}{1-\gamma} \qquad \text{with the portfolio separation property.}$$

We therefore assume for induction a similar form for J_t. Putting $c_{t-1} = k_{t-1} x$, the induction step becomes

$$J_{t-1}(x, \mathbf{v}) = x^{1-\gamma} \sup_{k,\mathbf{u}} \{\frac{A_{t-1} k^{1-\gamma} + (1-k)^{1-\gamma} \mathsf{E}_{t-1}\big[K_t(\mathbf{V}_t)(1+r+\mathbf{u}\mathbf{V}_t)^{1-\gamma}\big]}{1-\gamma}\}$$

where again, the optimal values are independent of x, establishing the portfolio separation property (and the new K_{t-1} is ≥ 0 since K_t is). We are done with the proof. Any time-dependence of $\mathbf{V}_{t+1}|\mathbf{V}_t$ will affect the E_{t+1} operator and hence the value of the K_t coefficient, but not the form.

(b) Infinite-horizon case; note that the payoffs are either always positive or always negative, so Seierstad conditions (P) or (N) hold, and the Bellman equation (which is based on time-invariance, in particular of $\mathbf{V}_{t+1}|\mathbf{V}_t$) is sufficient. It is natural to guess a form $K(\mathbf{v})x^{1-\gamma}/(1-\gamma)$ here as well, and $c_t = k_t X_t$, in order to try to apply sufficient conditions. Inserting this into the Bellman equation

$$J(x,\mathbf{v}) = \sup_{k,\mathbf{u}}\{\frac{(kx)^{1-\gamma}}{1-\gamma} + \alpha \mathsf{E}_{\mathbf{v}}\big[K(\mathbf{V})\frac{\big(x(1-k)(1+r+\mathbf{u}\mathbf{V})\big)^{1-\gamma}}{1-\gamma}\big]\}$$

we get

$$J(x,\mathbf{v}) = x^{1-\gamma}\sup_{k,\mathbf{u}}\{\frac{k^{1-\gamma} + \alpha(1-k)^{1-\gamma}\mathsf{E}_{\mathbf{v}}\big[K(\mathbf{V})(1+r+\mathbf{u}\mathbf{V})^{1-\gamma}\big]}{1-\gamma}\}$$

Now the optimization wrt. $\mathbf{u}$ separates out, giving a $\mathbf{u}^*$ not depending on x (and hence two fund monetary separation), while for $K > 0$, we have a concave problem to find the optimal $k^* = k^*(\mathbf{v})$ (which does not depend on x), whose solution is given by the first-order condition

$$0 = (k^*)^{-\gamma} - \alpha\eta(\mathbf{v})(1-k^*)^{-\gamma},$$

$$\text{i.e.} \qquad k^* = \Big(1 + \big[\alpha\eta(\mathbf{v})\big]^{1/\gamma}\Big)^{-1}$$

where η is the expectation term. We get $K(\mathbf{v}) = (k^*)^{-\gamma}$, where k^* depends on the *function* K – potentially all its values – through η, and in order to apply sufficiency it actually remains to show that there exists a positive function $K(\mathbf{v})$ satisfying this functional equation. This is probably beyond what is expected in a course like ECON5160. (However, it is not hard to show in the case where $\mathbf{V}$ is independent of $\mathbf{v}$, where K is not a function of $\mathbf{v}$ but merely a constant – then one may in fact find an explicit expression for K. We skip the details.)

(c) The bound is zero, so as long as the Bellman equation holds, it is no issue for the (P) or (N) conditions whether α were 1.

(d) This question concerns the restrictions to nonnegative portfolio weights, which has the interpretations of forbidding short sales.
Inspecting the calculations, we see that the maximizations which were independent of x still are. Hence all properties carry over; the optimized program may of course be less favourable since the opportunity set is restricted, but that affects only the actual choices and the K values, not the form.

 # 3 Dynamics: $X_{t+1} = (X_t - c_t)(1 + r + \mathbf{u}_t\mathbf{V}_{t+1})$ as in problem #2.

(a) Again, the calculations will not assume $\mathbf{V}_{t+1}|\mathbf{V}_t$ time-invariant. The «updated» problem was slightly more general,

$$\sup_{\mathbf{u},c} \sum_{t=0}^{T-1} -A_t e^{-a_t c_t} - e^{-a_T X_T}$$

We then have

$$\begin{aligned} J_T(x, \mathbf{v}) &= -e^{-a_T x} \\ J_{T-1}(x, \mathbf{v}) &= \sup_{\mathbf{u},c}\{-A_{T-1}e^{-a_{T-1}\cdot c} - \mathsf{E}_{T-1}[\exp(-a_T(x-c)(1+r+\mathbf{u}\mathbf{V}_T))]\} \\ &= \sup_{\mathbf{u},c}\{-A_{T-1}e^{-a_{T-1}\cdot c} - e^{-a_T(x-c)(1+r)}\mathsf{E}_{T-1}[\exp(-a_T(x-c)(\mathbf{u}\mathbf{V}_T))]\} \end{aligned}$$

Writing $(x - c)\mathbf{u} = \mathbf{b}$, the optimization wrt. $\mathbf{u}$ becomes $\inf_{\mathbf{b}} \mathsf{E}_{T-1}[\exp(-a_T\mathbf{b}\mathbf{V}_T)]$, which does not depend on x nor c. Assuming an optimal $\mathbf{b}*$ exists, we have two-fund monetary separation. With $\eta_{T-1}(\mathbf{v}) = \mathsf{E}_{T-1}[\exp(-a_T\mathbf{b}^*\mathbf{V}_T)]$, we have for the optimization wrt. c:

$$J_{T-1}(x, \mathbf{v}) = \sup_{c}\{-A_{T-1}\exp(-a_{T-1}\cdot c) - \eta_{T-1}(\mathbf{v})\exp(-a_T(x-c)(1+r))\}$$

where the braced expression is concave in c. The control region is an interval $[0, x]$ so either the optimal c^* is a stationary point or an endpoint, and in the latter case, we will have J of the form

$$-R_{T-1}(\mathbf{v}) - Q_{T-1}(\mathbf{v})\exp(-q_{T-1}x)$$

(all coefficients positive, and at most one depends on $\mathbf{v}$). In case of a stationary point, we have

$$a_{T-1}A_{T-1}\exp(-a_{T-1}\cdot c) = a_T(1+r)\eta_{T-1}(\mathbf{v})\exp(-a_T(x-c)(1+r))$$

so that c^* is linear in x, and

$$\begin{aligned} J_{T-1}(x, \mathbf{v}) &= -A_{T-1}(1 + \frac{a_{T-1}}{a_T(1+r)})\exp(-a_{T-1}\cdot c), \qquad \text{which is of the form} \\ &= -Q_{T-1}\exp(-q_{T-1}x) \end{aligned}$$

So for induction, we assume the most general form $J_t(x, \mathbf{v}) = -R_t(\mathbf{v}) - Q_t(\mathbf{v})\exp(-q_t x)$. Then for J_{t-1}, we get

$$\begin{aligned} J_{t-1}(x, \mathbf{v}) &= \sup_{\mathbf{u},c}\{-A_{t-1}e^{-a_{t-1}\cdot c} - \mathsf{E}_{t-1}[Q_t(\mathbf{V}_t)\exp(-q_t(x-c)(1+r+\mathbf{u}\mathbf{V}_t)) + R_t(\mathbf{V}_t)]\} \\ &= \sup_{\mathbf{u},c}\{-A_{t-1}e^{-a_{t-1}\cdot c} - e^{-a_t(x-c)(1+r)}\mathsf{E}_{t-1}[Q_t(\mathbf{V}_t)\exp(-q_t(x-c)(\mathbf{u}\mathbf{V}_T))]\} \\ &\quad - \mathsf{E}_{t-1}[R_t(\mathbf{V}_t)] \end{aligned}$$

where again, the maximization with respect to $\mathbf{b} = (x-c)\mathbf{u}$ yields two-fund monetary separation, and

$$J_{t-1}(x,\mathbf{v}) = \sup_c\{-A_{t-1}\exp(-a_{t-1}\cdot c) - \eta_{t-1}(\mathbf{v})\exp(-q_t(x-c)(1+r))\} - \mathsf{E}_{t-1}[R_t(\mathbf{V}_t)]$$

where the braced expression has same form as in the calculations for $T-1$, except with q_t rather than a_t in the second exponential. We omit the details.

(b) Answer as in 2d.

(c) As always, increased opportunity set might increase the value function, but here we are interested in the form. First a couple of comments on the material content of increasing the opportunity set: Allowing for $c > x$ means we are allowed to borrow in the market, and will be penalized accordingly if we have negative wealth in the end – which makes sense as our utility function is defined everywhere for all parameters, in contrast to the CRRA case. Allowing for negative c might seem a bit nonsensical from an applied point of view though; it might be interpreted as borrowing from a non-market source (which would rather have been used for consumption) in order to invest in the market, but this source then has to be of arbitrary size. However from a mathematical point of view, the problem makes perfect sense.
In mathematical terms, the new control regions mean that one or both of the endpoint solutions for c will be ruled out. Consider

$$\sup_c\{-A_{t-1}\exp(-a_{t-1}\cdot c) - \eta_{t-1}(\mathbf{v})\exp(-q_t(x-c)(1+r))\} - \mathsf{E}_{t-1}[R_t(\mathbf{V}_t)].$$

The positiveness of the coefficients means that there actually *is* a stationary point in $(-\infty,\infty)$, so relaxing the control region does not affect existence of solution. It will turn out that the form $J_t(x,\mathbf{v}) = -R_t(\mathbf{v}) - Q_t(\mathbf{v})\exp(-q_t x)$ which was in the general case necessary to include for the case $c \in [0,x]$, still holds. This is the essential answer to this problem; Feinschmeckers might want to add that some parameter(s) will become independent of $\mathbf{v}$ or even vanish, so the control region $[0,x]$ does in fact require the most generality of the four considered. For completeness, we give the arguments:

- Consider first the case $(-\infty,\infty)$, we where the only candidate is the (unique) stationary point, which yields the form

  $$-\mathsf{E}_{t-1}[R_t(\mathbf{V}_t)] - Q_{t-1}\exp(-q_{t-1}x)$$

  – which is no more general than before.
  «Feinschmecking»: it follows by easy backwards induction that all $R_t = 0$ and the Q_t do not depend on $\mathbf{v}$.
- For the case $(-\infty,x]$, we either have an internal solution as in the previous point, or $c = x$ with the form $-\mathsf{E}_{t-1}[R_t(\mathbf{V}_t)] - \eta_{t-1}(\mathbf{v}) - Q_{t-1}\exp(-q_{t-1}x)$, which is no more general than before.

«Feinschmecking»: by induction starting from Q_T, no Q_t will depend on $\mathbf{v}$ – however, there will be a $\mathbf{v}$-dependent R term (except if by coincidence all optima are stationary points).

- For the case $[0, \infty)$, we will if $c^* = 0$, have $Q_t(\mathbf{v}) = \eta_t(\mathbf{v})$, and $R_{t-1} = A_{t-1} + \mathsf{E}_{t-1}[R_t(\mathbf{V}_t)]$, so the form is no more general than before.
  «Feinschmecking»: Inspecting the forms for R_{t-1} for the boundary solution and for the internal solution, it will however follow (induction from T again), that no R_t depends on $\mathbf{v}$.

Problems from TK chapter V

TK V P. 3.4

$$\text{For } w_1: \qquad \int_{w_1}^{\infty} \lambda^2 e^{-\lambda w_2}\,\mathrm{d}w_2 = \lambda\Big[e^{-\lambda w_2}\Big]_{\infty}^{w_1} = \lambda e^{-\lambda w_1}$$
$$\text{For } w_1: \qquad \int_0^{w_2} \lambda^2 e^{-\lambda w_2}\,\mathrm{d}w_1 = \lambda^2 w_2 e^{-\lambda w_2}$$

as should be.

TK V P. 4.2 See the example on TK p. 301 ff.; In the problem, $X(t)$ has the role of $M(t)$ in the example, and $N = X+Y$ the role of X. So $\mathsf{Pr}[X(t) = m | N(t) = n]$ has the binomial distribution given on top of p. 303 with p given by (4.5) p. 302; then the joint probability

$$\begin{aligned}
\mathsf{Pr}[(X(t), Y(t)) = (x, y)] &= \mathsf{Pr}[(X(t), N(t)) = (x, x+y)] \\
&= \mathsf{Pr}[N(t) = x+y]\mathsf{Pr}[X(t) = x | N(t) = x+y] \\
&= \frac{(\lambda t)^{x+y}}{(x+y)!} e^{-\lambda(x+y)} \frac{(x+y)!}{x!y!} p^x (1-p)^y \\
&= \frac{(\lambda t)^{x+y}}{x!y!} e^{-\lambda(x+y)} \Big\{\frac{1}{t}\int_0^t (1-G(z))\,\mathrm{d}z\Big\}^x \Big\{1 - \frac{1}{t}\int_0^t (1-G(z))\,\mathrm{d}z\Big\}^y \\
&= \frac{\lambda^{x+y}}{x!y!} e^{-\lambda(x+y)} \Big\{\int_0^t (1-G(z))\,\mathrm{d}z\Big\}^x \Big\{\int_0^t G(z)\,\mathrm{d}z\Big\}^y.
\end{aligned}$$

TK V P. 4.7 Should be near-trivial after the stochastic integral wrt. $\mathrm{d}N$ is introduced.

TK V P. 6.3 Consider the examples in TK section 6.1; we will have $\lambda^{-1}\sum_{n=0}^{\infty} G^{(n)}(a-1)$, analogous to the formula middle of p. 320, but with $a-1$ because the state space is the nonnegative integers, and state $= a$ means failure, and with $G^{(n)}$ being *discrete* n-fold convolution, defined inductively by

$$\begin{aligned}
G^{(0)}(x) &= 1_{x\ge 0}, \qquad \text{and for } n > 0 \\
G^{(n)}(x) &= \mathsf{Pr}[Y_1 + \cdots + Y_n \le x] \\
&= \sum_{y=0}^{x} \mathsf{Pr}[Y_1 + \cdots + Y_{n-1} \le x-y]\mathsf{Pr}[Y_n = y] \\
&= \sum_{y=0}^{x} G^{(n-1)}(x-y)\mathsf{Pr}[Y_n = y] \\
&= \sum_{y=0}^{x} G^{(n-1)}(y)\mathsf{Pr}[Y_n = x-y],
\end{aligned}$$

where in this case we have an explicit distribution for Y; denote its cdf (equalling $G^{(1)}$) by F. Going from here requires a bit of a trick[2]: Put $b = a - 1$ for notational convenience, and assume $b \geq 0$ so that $1_{b\geq 0} = 1$. Denoting $\sum_{n=0}^{\infty} G^{(n)}(x)$ by $v(x)$, we have

$$\begin{aligned}
\lambda \mathsf{E}[T] = v(b) &= \sum_{n=0}^{\infty} G^{(n)}(b) \\
&= 1_{b\geq 0} + \sum_{n=1}^{\infty}\sum_{y=0}^{b} G^{(n-1)}(y)\mathsf{Pr}[Y_n = b-y] \\
&= 1 + \sum_{y=0}^{b}\sum_{n=1}^{\infty} G^{(n-1)}(y)\mathsf{Pr}[Y = b-y] \\
&= 1 + \sum_{y=0}^{b} v(y)\mathsf{Pr}[Y = b-y]
\end{aligned}$$

(by shifting summation index by one,) so that

$$\lambda \mathsf{E}[T] = v(b) = 1 + \sum_{y=0}^{b} v(y)\mathsf{Pr}[Y = b-y]$$

which gives a recursion formula for v:

$$v(b)(1 - \mathsf{Pr}[Y = 0]) = 1 + \sum_{y=0}^{b-1} v(y)\mathsf{Pr}[Y = b-y].$$

This far we have not used any specific form for the distribution of Y, only that it lives on the nonnegative integers. Inserting for the geometric distribution

$$\begin{aligned}
v(b)(1-p) &= 1 + \sum_{y=0}^{b-1} v(y)p(1-p)^{b-y} \\
&= 1 + \underbrace{\sum_{y=0}^{b-2} v(y)p(1-p)^{b-1-y}}_{=v(b-1)(1-p)-1}(1-p) + v(b-1)p(1-p) \\
&= 1 - (1-p) + v(b-1)(1-p)\cdot(1-p+p) \\
&= p + (1-p)v(b-1)
\end{aligned}$$

– a first-order difference equation $v(x+1) = \frac{p}{1-p} + v(x)$ with solution

$$v(b) = b\frac{p}{1-p} + v(0)$$

[2] «trick» as in «tricky», I'd add.

(doesn't this make you suspect it could have been done simpler?) We need to determine

$$v(0) = \sum_{n=0}^{\infty} G^{(n)}(0) = 1 + \sum_{n=1}^{\infty} \mathsf{Pr}[Y_1 = Y_2 + \cdots + Y_n = 0] = 1 + \sum_{n=1}^{\infty} p^n = (1-p)^{-1}$$

where we used that the sum of the Y_n is zero if and only if all Y_n are. So, finally,

$$\mathsf{E}[T] = \frac{1}{\lambda} v(b) = \frac{1}{\lambda}[\frac{bp}{1-p} + \frac{1}{1-p}] = \lambda^{-1}(1 + \frac{ap}{1-p})$$

– with lots of reservations for possible errors.

TK V P. 6.10 *(This one was certainly difficult; the below approach is ... well, doable, but if anyone has made it through the conditional uniform distribution given $N(1) = n$, please let me know. Oh, and if you got a different answer, you might be right. Or we might both be wrong.)*

I will first calculate as far as possible without any assumption on the acceptance threshold $\theta(t)$, which may afterwards be put equal to a constant ϑ or a function as specified.

You reject bids below a certain threshold $\theta(t)$, and since bids are uniformly distributed, that means that a bid occuring at time t has probability $1-\theta(t)$ of being accepted. Assume now that you only see high enough bid – too low bids are rejected by an auctioneer. Then the process $X(t)$ of bids arriving to your knowledge, is a marked inhomogeneous Poisson process with time-dependent intensity $\lambda(t) = 1 - \theta(t)$ (since the original Poisson process has intensity 1); the number of bids you see at time t, is distributed Poisson with parameter $\Lambda(t) = \int_0^t \lambda(s)\,\mathrm{d}s$. Let T be the time of the first sufficiently high bid, i.e. the winning one; then

$$\mathsf{Pr}[T > t] = \mathsf{Pr}[\#\text{bids} = 0] = \exp(-\Lambda(t)), \quad \text{and in particular}$$
$$\mathsf{Pr}[\text{sold at time } 1] = \mathsf{Pr}[T \leq 1] = 1 - \exp(-\Lambda(1))$$

and also that T has pdf $= \lambda(t)\exp(-\Lambda(t))$. Now let V be the size of the winning bid; conditional on T, it is uniform on $[1-\lambda(t), 1]$ if $T \leq 1$ and 0 otherwise, and has conditional expectation $(1 - \frac{1}{2}\lambda(T)) \cdot 1_{T \leq 1}$. The double expectation law gives

$$\begin{aligned}\mathsf{E}[V] = \mathsf{E}[\mathsf{E}[V|T]] &= \mathsf{E}[(1 - \frac{1}{2}\lambda(T)) \cdot 1_{T\leq 1}] \\ &= \mathsf{Pr}[T \leq 1] - \frac{1}{2}\int_0^1 (\lambda(t))^2 e^{-\Lambda(t)}\,\mathrm{d}t.\end{aligned}$$

Time to insert. For constant θ, i.e. $\lambda = 1 - \vartheta$, we have

$$\mathsf{Pr}[T \leq 1] = 1 - \exp(\vartheta - 1) \qquad \text{and}$$
$$\begin{aligned}\mathsf{E}[V] &= 1 - \exp(\vartheta - 1) - \frac{1}{2}(\vartheta - 1)^2 \int_0^1 \exp((\vartheta - 1)t)\,\mathrm{d}t \\ &= \frac{1}{2}(\vartheta + 1)(1 - \exp(\vartheta - 1))\end{aligned}$$

so that the maximizing ϑ should[3] satisfy the first-order condition $1 = (\vartheta + 2)\exp(\vartheta - 1)$.

For the given function $\theta(t) = (1-t)/(3-t)$, which gives $\lambda(t) = 2/(3-t)$ and $\exp(-\Lambda(t)) = (1 - t/3)^2$, we have

$$\begin{aligned}
\mathsf{Pr}[T \le 1] &= 1 - \exp(-\Lambda(1)) = 1 - (1 - 1/3)^2 = \frac{5}{9}, \qquad \text{and} \\
\mathsf{E}[V] &= 5/9 - \frac{1}{2}\int_0^1 4(\frac{1 - t/3}{3 - t})^2 \, \mathrm{d}t \\
&= \frac{1}{3}
\end{aligned}$$

(Numerical calculations show that this is a small improvement – the best constant ϑ gives a payoff < 0.3305 for $\vartheta \approx 0.21$.)

[3] ...objections, anyone? I would be inclined to guess that a textbook problem like this should admit a more explicit answer?

Problems for March 20

TK VI P. 4.1 Consider the repairman problem is given on p. 369 ff. The infinitesimal generator of the process is

$$\mathbf{A} = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} \begin{array}{c} \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\ \left[\begin{array}{cccccc} -\lambda R & \lambda R & 0 & 0 & 0 & 0 \\ \mu & -(\mu+\lambda R) & \lambda R & 0 & 0 & 0 \\ 0 & 2\mu & -(2\mu+\lambda R) & \lambda R & 0 & 0 \\ 0 & 0 & 3\mu & -(3\mu+\lambda R) & \lambda R & 0 \\ 0 & 0 & 0 & 4\mu & -(4\mu+\lambda R) & \lambda R \\ 0 & 0 & 0 & 0 & 5\mu & -5\mu \end{array}\right] \end{array}$$

Note that Gaussian elimination for an equation system $\boldsymbol{\pi}\mathbf{A} = \mathbf{0}$, is performed by *column* operations. In this case, it is particularly easy; add the last column ($\#$ 5) to $\#$ 4; the new $\#$ 4 to $\#$ 3 etc. (and if you start writing from the last column, you don't need to write more than one more matrix). Then column $\#$ 0 can be deleted, π_0 chosen freely, solve and normalize, and insert in the formulas p. 370 and top p. 371.

TK VI P. 6.1 (Is the Markov argument something you want to see with a formal proof?) The infinitesimal parameter λ_{yz} for state y to state z, is λP_{yz} unless $y = z$, for which it is $-\lambda \sum_{z \neq y} P_{yz}$.

TK VII P. 3.1 By independence, we can evaluate separately and then multiply. However, having not done exercise 3.3, the easiest way to evaluate $\mathsf{E}[W_{N(t)+1}]$ is maybe to condition on $N(t)$ nevertheless: $\mathsf{E}[W_{N(t)+1}|N(t)+1 = n+1]$ is the $n+1$st waiting time, with a mean of $(n+1)/\lambda$. So $\mathsf{E}[\mathsf{E}[W_{N(t)+1}|N(t)+1]] = \mathsf{E}[(N(t)+1)/\lambda] = t + \frac{1}{\lambda}$. For $\mathsf{E}[(N(t)+1)^{-1}]$, we just perform the calculations:

$$\mathsf{E}[(N(t)+1)^{-1}] = \sum_{n=0}^{\infty} \frac{1}{n+1} \frac{(\lambda t)^n}{n!} e^{-\lambda t} = \sum_{n=0}^{\infty} \frac{1}{\lambda t} \frac{(\lambda t)^{n+1}}{(n+1)!} e^{-\lambda t} = \frac{1}{\lambda t}(1 - e^{-\lambda t})$$

Multiplying the factors, we get the result.

Problems for March 27

Note, P. 1.1

(a) Let $X(t) = \sigma_1 B_1(t) + \cdots + \sigma_n B_n(t)$. The increments $X(t+h) - X(t)$ are independent for disjoint time intervals (since so is the case for the B_i) and Gaussian with mean 0 and variance $\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}$ (since B_i, B_j are independent for $i \neq j$). So it is necessary and sufficient that the vector norm $\boldsymbol{\sigma} \cdot \boldsymbol{\sigma}$ is $= 1$.

(b) Define $Y(t) = \int_0^t \boldsymbol{\sigma} \cdot \mathrm{d}\mathbf{B}(s)$. Then similar to (a), Y has increments which are independent of the past, and Gaussian with mean zero and variance $\int_t^{t+h} \boldsymbol{\sigma}(s) \cdot \boldsymbol{\sigma}(s)\,\mathrm{d}t$. For this to be equal to h, then necessarily its h-derivative – equalling the integrand evaluated at t – must be 1.[4] We easily see that $\boldsymbol{\sigma}(t) \cdot \boldsymbol{\sigma}(t) = 1$ for all t is also sufficient.

(c) Write $\boldsymbol{\sigma}(t) = \boldsymbol{\sigma}(t, \varpi)$ where t is time and ϖ is chance[5]; similarly, write $\mathbf{B}$ as $\mathbf{B}(t, \omega)$. Conditional on ϖ – i.e., knowing the entire time series $\{\boldsymbol{\sigma}\}_{t \geq 0}$ – then by independence, $\mathbf{B}$ is still a standard Brownian vector, and $Z(t) = \int_0^t \boldsymbol{\sigma}(s, \varpi) \cdot \mathrm{d}\mathbf{B}(s, \omega)$ is a well-defined Itô integral *for each* ϖ, with a conditionally non-random integrand, hence a time-inhomogeneous (not standard) Brownian motion whose increments have variance $\int_t^{t+h} \boldsymbol{\sigma}(s, \varpi) \cdot \boldsymbol{\sigma}(s, \varpi)\,\mathrm{d}s$. So Z is unconditionally Gaussian. Now the increments have t-conditional variance given by the law of total variance[6]; since the increments have zero mean, the t-conditional variance may be obtained by first conditioning on ϖ. We obtain $\int_t^{t+h} \mathsf{E}_t[\boldsymbol{\sigma}(s, \varpi) \cdot \boldsymbol{\sigma}(s, \varpi)]\,\mathrm{d}s$; just as in (b), it turns out both necessary and sufficient that $\mathsf{E}_t[\boldsymbol{\sigma}(t, \varpi) \cdot \boldsymbol{\sigma}(t, \varpi)] = 1$ for all t.[7]

Corr. 04–02

Note, P. 1.2 Actually, the assumption in 1.3 of continuous X should have been stated here too – otherwise, we should use t^- instead of t in the integrands. Also the first line in (a) lacked a «Show that» at the end.

(a) Inserting for $\mathrm{d}Z_i = Z_i\,\mathrm{d}X_i$ in the self-financingness definition, we have

$$\mathrm{d}Y(t) = \nu_0(t) Z_0(t)\,\mathrm{d}X_0(t) + \nu_1(t) Z_1(t)\,\mathrm{d}X_1(t) + \cdots + \nu_n(t) Z_n(t)\,\mathrm{d}X_n(t)$$

where of course $\nu_i Z_i$ equals quantum times price, equals value v_i of your position in opportunity number i. They have to sum up to Y, so inserting for the v_i and for $v_0 = Y - \sum_{i=1}^n v_i$,we get

$$\mathrm{d}Y(t) = [Y(t) - \sum_{i=1}^n v_i(t)]\,\mathrm{d}X_0(t)) + \sum_{i=1}^n v_i(t)\,\mathrm{d}X_i(t).$$

Note that $\sum_{i=1}^n v_i = \mathbf{v} \cdot \mathbf{1}$, and the conclusion follows.

[4] At least for «almost every» t, but that detail is not essential in this course.

[5] in mathematical terms, the element of the underlying probability space

[6] see e.g. http://en.wikipedia.org/w/index.php?title=Law_of_total_variance&oldid=273626546

[7] Actually, it is possible to read off $\boldsymbol{\sigma}$ up to time t from observing $Z(\to t)$; heuristically, $(\mathrm{d}Z)^2/\mathrm{d}t = \boldsymbol{\sigma} \cdot \boldsymbol{\sigma}$ has to equal 1, but the proof that $\lim_h |Z(t+h) - Z(t)|^2/|h| = \boldsymbol{\sigma} \cdot \boldsymbol{\sigma}$ is beyond what is taught.

(b) Since Z_0 is only driven by a $\mathrm{d}t$-integral, the Itô formula gives no second-order term:

$$\begin{aligned}\mathrm{d}\frac{Y}{Z_0} &= \frac{\mathrm{d}Y}{Z_0} - \frac{Y}{Z_0^2}\,\mathrm{d}Z_0 \\ &= \frac{1}{Z_0}[Y\,\mathrm{d}X_0 + \mathbf{v}\cdot(\,\mathrm{d}\mathbf{X} - \mathbf{1}\,\mathrm{d}X_0)] - \frac{Y}{Z_0}\frac{\mathrm{d}Z_0}{Z_0} \\ &= \frac{\mathbf{v}}{Z_0}\cdot\,\mathrm{d}\hat{\mathbf{X}}\end{aligned}$$

since $\mathrm{d}Z_0 = Z_0\,\mathrm{d}X_0$.

Note, P. 1.3

(a) Because there are two mixed second-order derivatives, $f''_{xy} = 1 = f''_{yx}$, the sum of 2 cancelling the $\frac{1}{2}$.

(b) $\mathrm{d}(B^2 - t) = 2B\,\mathrm{d}B + \frac{1}{2}\cdot 2(\,\mathrm{d}B)^2 - \,\mathrm{d}t = 2B\,\mathrm{d}B$. Also, $\mathsf{E}[|B^2 - t|] < \infty$.

(c) We differentiate to get $\mathrm{d}(F(B) - Z) = F'(B)\,\mathrm{d}B + \frac{1}{2}F''(B)\,\mathrm{d}t - \,\mathrm{d}Z$, which – modulo the integrability condition – is a martingale if $Y(t) = \frac{1}{2}F''(B(t))$.

(d) i. The absolute value is a convex function, and Jensen's inequality may be employed. But you might even simpler argue that $\mathsf{E}_t|B(t+h)| \geq \mathsf{E}_t[B(t+h)\text{ sign }B(t)] = B(t)\text{ sign }B(t) = |B(t)|$.

ii. Ouch, yet another typo: Should have been «where A does not hit 0».
Either $B(t) > 0$ on the interval, or $B(t) < 0$ on the interval. In the first case, $A = B$ so $\mathrm{d}A = \,\mathrm{d}B$, and in the latter we have $\mathrm{d}A = -\,\mathrm{d}B$. So $U = \text{sign }B$ works; this is undefined when $B = 0$, but we can choose an arbitrary definition then.

iii. $U\,\mathrm{d}B = \,\mathrm{d}A$ except when $A = 0$, so we have $\mathrm{d}Z = 1_{A=0}(\,\mathrm{d}A + U\,\mathrm{d}B)$; since we are free to define U when $B = 0$, we define it to be zero then. So $\mathrm{d}Z = 1_{A=0}\,\mathrm{d}A$. This shows that Z only changes when $A = 0$. By our interpretation of the stochastic integral, «$\int_0^t 1_{A(s)=0}\,\mathrm{d}A$» is equal to $\lim_{h\to 0^+}\int_0^t 1_{A(s)=0}\frac{A(s+h)-A(s)}{h}\,\mathrm{d}s$, and the fraction is positive[8] since $A(s) = 0 \leq A(s+h)$.
Note: Such Z, which are non-constant, nondecreasing, and increasing on a set of zero total time, show up in e.g. «consume so that wealth does not exceed a a given threshold» models.

(e) The integral is $\frac{1}{2}(X(T) - X(0))^2$, which is of course nonnegative.
(Check: differentiate to get $(X(t) - X(0))\,\mathrm{d}(X(t) - X(0))$, where $X(0)$ is a constant.)

[8]You think this is a bit insufficient? For full rigor, it is certainly too sloppy yes.

(f) Since $\mathrm{d}X(t) = \mathrm{d}(X(t) - X(0))$, we can put $Y(t) = X(t) - X(0)$ and integrate by parts to get

$$\int_0^T \max\{0, Y(t)\}\,\mathrm{d}Y(t) = \big[\max\{0, Y(T)\}Y(T)\big]_0^T - \int_0^T Y(t)(1_{Y(t)>0}\,\mathrm{d}Y(t))$$
$$= \max\{0, Y(T)\}^2 - \int_0^T \max\{0, Y(t)\}\,\mathrm{d}Y(t)$$

so that $\int_0^T \max\{0, X(t) - X(0)\}\,\mathrm{d}X(t) = \frac{1}{2}\max\{0, X(T) - X(0)\}^2$.

(g) In both cases, the gain from trading is nonnegative. Case (i) gives a free lunch except if $X(T) = X(0)$, so unless X is constant, we can find a T for which $\mathsf{Pr}[X(T) \neq X(0)] > 0$, with a riskless chance of a positive return. Case (i) gives a free lunch if $X(T) > X(0)$, and gives the same result unless X is nonincreasing – in which case you can *short-sell* the same portfolio and have a riskless free lunch. (Of course, in the case where X is nonincreasing and shorting is forbidden, you never buy for trading!)
Bottom line: we need some other change of variables rule.

Note, P. 1.4

(a) For the continuous part, use the calculations from 1.2. For the slightly complicating $\mathrm{d}N$ contribution, it is actually fully rigorous to just observe that at a jump time, Z jumps by a *factor* of $\gamma(t^-)$ while Z_0 does not jump.

(b) You should be able to take care of the continuous part by 1.3 (a), where $(\sigma\,\mathrm{d}B))(-\theta\,\mathrm{d}B) = -\sigma\theta\,\mathrm{d}t$. For the $\mathrm{d}N$ contribution, consider what happens at a jump time: $\hat{Z}$ jumps by a factor γ, G jumps by a factor η, so that the post-jump state divided by the pre-jump state, equals $(1+\gamma)(1+\eta)$; then the *change* is by a factor of $(1+\gamma)(1+\eta) - 1$.

(c) The $\mathrm{d}B$ term is already a martingale, while

$$(\eta + \gamma + \eta\gamma)\,\mathrm{d}N = (\eta + \gamma + \eta\gamma)\,\mathrm{d}\tilde{N} + (\eta + \gamma + \eta\gamma)\lambda\,\mathrm{d}t$$

where the $\mathrm{d}\tilde{N}$ term is martingale. Deleting the martingale parts and equating the remaining terms to 0, we get

$$0 = \mu - r - \sigma\theta + \eta\gamma\lambda. \qquad («\theta»)$$

(d) Arrgh, congratulations Nils with yet another typo. This time, it was intended $\gamma = \eta = 0$ – or equivalently, $\lambda = \eta = 0$ – but only $\eta = 0$ was written.
So, let me do the differentiation without any assumptions first, and then comment. In the following, «(t)» arguments are suppressed, but «(t^-)» are not:

$$\begin{aligned}\mathrm{d}(GF(\hat{Z})) = {} & GF'(\hat{Z})\hat{Z} \cdot ((\mu - r - \lambda\gamma)\,\mathrm{d}t + \sigma\,\mathrm{d}B) \\ & + F(\hat{Z})G \cdot (-\lambda\eta\,\mathrm{d}t - \theta\,\mathrm{d}B) \\ & - \theta GF'(\hat{Z})\hat{Z}\sigma\,\mathrm{d}t \\ & + \{G(t^-)(1+\eta(t^-))F([1+\gamma(t^-)]\hat{Z}(t^-)) - G(t^-)F(\hat{Z}(t^-))\}\,\mathrm{d}N\end{aligned}$$

which, by writing $\mathrm{d}N = \mathrm{d}\tilde{N} + \lambda\,\mathrm{d}t$, equals

$$\begin{aligned}
&= GF'(\hat{Z})\hat{Z}\cdot((\mu - r - \sigma\theta - \lambda\gamma)\,\mathrm{d}t + \\
&\quad + \{F((1+\gamma)\hat{Z}) - F(\hat{Z})\}(1+\eta)G\lambda\,\mathrm{d}t \\
&\quad + [\text{« d}B \text{ and d}\tilde{N}\text{» terms}]
\end{aligned}$$

Now, with the θ_0 set as to satisfy formula («θ»), we get

$$\begin{aligned}
&= \{F((1+\gamma)\hat{Z}) - F(\hat{Z}) - \gamma\hat{Z}F'(\hat{Z})\}(1+\eta)G\lambda\,\mathrm{d}t \\
&\quad + [\text{« d}B \text{ and d}\tilde{N}\text{» terms}]
\end{aligned}$$

(Observe that the first-order approximation of the braced expression wrt. γ, is zero.)

~~Now, I re-assign the problem for you: If $\lambda = 0$, what then?~~
With $\lambda = 0$, the $\mathrm{d}t$ term vanishes and we are left with a martingale. So this new probability measure defined by $\mathsf{E}_Q[X] = \mathsf{E}[XG]$ does not only make $\hat{Z}$ a martingale, but also $F(\hat{Z})$. *Added 04–01*

(e) ~~Re-assign the problem with the corrected condition $\lambda = 0$.~~
$\mathsf{E}_t[G(T)F(\hat{Z}(\tau))] = \mathsf{E}_t[G(\tau)F(\hat{Z}(\tau))]$ since G is a martingale. By optional sampling, it doesn't matter that there is a τ argument and not a deterministic time in the latter formula, so we are left with $G(t)F(\hat{Z}(t))$.[9] *Added 04–01*

Note, P. 1.5

(a) A_Y applied on a function f, will yield $\mathsf{A}_Y f(t,z) = f'_t + z\mu(t)f'_z + \frac{1}{2}z^2(\sigma(t))^2 f''_{zz}$.
So with this $f = H$, we have

$$\mathsf{A}_Y H(t,z) = -\delta(t)H(t,z) + q\mu(t)H(t,z) + \frac{1}{2}q(q-1)(\sigma(t))^2 H(t,z)$$

which vanishes if

$$\delta(t) = q\cdot(\mu(t) + \frac{1}{2}(q-1)(\sigma(t))^2.$$

(b) With F as hinted, we have $\mathsf{A}_Y F(t,z) = -rF(t,z) + q\mu F(t,z) + \frac{1}{2}q(q-1)\sigma^2 F(t,z)$, so

$$\mathsf{A}_Y F(t,z) + z^q e^{-rt} = z^q e^{-rt}\cdot((-r + q\mu + \frac{1}{2}q(q-1)\sigma^2)A + 1)$$

so that $A = -(-r + q\mu + \frac{1}{2}q(q-1)\sigma^2)^{-1}$ if this is well-defined – if not, you will not be expected to solve it[10].

(c) $A_N f(x) = \lambda(f(x+1) - f(x))$.

[9] Comment: The random variable $G(T)$ works as a pricing kernel not only for T-claims (i.e. derivatives maturing at T), but also for τ-claims for any $\tau \le T$ – as long as we have a finite horizon, we need not more than one pricing kernel.

[10] It will be $Ae^{-rt}x^q\sum_{i=0}^{\infty} a_n x^n$, with the a_n defined by $a_0 = 1$ and $a_{n+1} = 2\frac{\sigma^2 q + \mu n}{n(n+1)\sigma^2}$.

Note, P. 2 (for April 3rd) OK, yet more typos and omissions :-(

- The solution requires a σ – required to be nonzero, in order to have a stochastic process – in front of the Itô integral
- The assumption that $\theta > 0$ – so that the process is really mean-reverting – was omitted by mistake.

The following will assume $\theta > 0 \neq \sigma$.

(a) I'll skip this – for the integrating factor approach, see part (e) below.

(b) Gaussianity follows from the argument of 1.1 (c), so we only need to establish limiting mean and variance. The Itô integral vanishes under expectation, and so the long-term mean converges to μ (here we have used $\theta > 0$), while the variance is

$$\sigma^2 e^{-2\theta t} \int_0^t e^{2\theta s}\,\mathrm{d}s = \frac{\sigma^2}{2\theta}(1 - e^{-2\theta t}) \qquad \text{which} \xrightarrow[t\to\infty]{} \frac{\sigma^2}{2\theta}$$

and the limiting distribution is $\mathcal{N}(\mu, \frac{\sigma^2}{2\theta})$.

(c) We have $\mathsf{E}[Z(t)] = \mu + e^{-\theta t}(\mathsf{E}[Y] - \mu)$, which converges to μ as in (b). For the variance:

$$\mathsf{var}\, Z(t) = e^{-2\theta t}\mathsf{E}[\big(Y - \mathsf{E}Y + \sigma \int_0^t e^{2\theta s}\,\mathrm{d}s\big)^2] = e^{-2\theta t}(\mathsf{var}\, Y + \sigma^2 \frac{e^{2\theta t} - 1}{2\theta})$$

by independence. It converges to $\frac{\sigma^2}{2\theta}$, as in (b).

(d) We must have $\mathsf{E}Y = \mu$ and $\mathsf{var}\, Y = \frac{\sigma^2}{2\theta}$. Furthermore, because Z is the sum of a Gaussian and $e^{-\theta t}(Y - \mathsf{E}[Y])$, and the latter tends to 0 in distribution, hence the limiting distribution is Gaussian, then Y has to be Gaussian too. We are done.

(e) We have $\mathrm{d}Z + \theta(t)Z\,\mathrm{d}t = \theta(t)\mu(t)\,\mathrm{d}t + \sigma(t)\,\mathrm{d}B(t)$. As integrating factor, we try $\exp(\Theta(t)) = \exp(\int_0^t \theta(s)\,\mathrm{d}s)$, so that

$$\mathrm{d}(e^{\Theta}Z) = e^{\Theta}(\,\mathrm{d}Z + \theta Z\,\mathrm{d}t) = e^{\Theta}\big(\theta\mu\,\mathrm{d}t + \sigma\,\mathrm{d}B\big)$$

i.e.
$$Z(t) = e^{-\int_0^t \theta(s)\,\mathrm{d}s}Z(0) + \int_0^t e^{-\int_s^t \theta(s')\,\mathrm{d}s'}\theta(s)\mu(s)\,\mathrm{d}s + \int_0^t e^{-\int_s^t \theta(s')\,\mathrm{d}s'}\sigma(s)\,\mathrm{d}B(s)$$

This is sufficent answer, but in order to compare it to the case with constant coefficients, one should maybe note that integration by parts in the $\mathrm{d}s$ integral, yields

$$\int_0^t e^{-\int_s^t \theta(s')\,\mathrm{d}s'}\theta(s)\mu(s)\,\mathrm{d}s = \mu(t) - \mu(0)e^{-\Theta(t)} - \int_0^t e^{-\int_s^t \theta(s')\,\mathrm{d}s'}\,\mathrm{d}\mu(s)$$

if μ is regular enough to permit it; this would then yield the form

$$Z(t) = \mu(t) + e^{-\int_0^t \theta(s)\,\mathrm{d}s}(Z(0) - \mu(0)) + \int_0^t e^{-\int_s^t \theta(s')\,\mathrm{d}s'}\,\mathrm{d}\mu(s) + \int_0^t e^{-\int_s^t \theta(s')\,\mathrm{d}s'}\sigma(s)\,\mathrm{d}B(s)$$

Notice the $\mathrm{d}\mu$ integral which vanishes if μ is constant.