Exam 2008 Problem 1

a) The survival function is given by

$$S(t) = P(T>t)$$

and the hazard rate is given by

$$\alpha(t) = \lim_{\Delta t \to 0} \frac{1}{\Delta t} P(t \leq T < t+\Delta t \mid T \geq t)$$

Now we have

$$\frac{1}{\Delta t} P(t \leq T < t+\Delta t \mid T \geq t)$$

$$= \frac{P(t \leq T \leq t+\Delta t)}{\Delta t} \cdot \frac{1}{P(T \geq t)}$$

$$= \frac{S(t) - S(t+\Delta t)}{\Delta t} \cdot \frac{1}{S(t)}$$

$$= - \frac{S(t+\Delta t) - S(t)}{\Delta t} \cdot \frac{1}{S(t)}$$

$$\to - S'(t) \frac{1}{S(t)} \quad \text{as } \Delta t \to 0$$

Thus we have

$$\alpha(t) = -\frac{S'(t)}{S(t)} = -\frac{d}{dt}\log S(t)$$

By integrating and using that $S(0) = 1$, we obtain

$$A(t) = \int_0^t \alpha(u)\,du = -\log S(t)$$

This gives

$$S(t) = e^{-A(t)} = e^{-\int_0^t \alpha(u)\,du}$$

We now let $T_1, T_2, \ldots, T_n$ be iid with survival function $S(t)$ and hazard rate $\alpha(t)$. We only observe the right censored survival times $\tilde{T}_1, \dots, \tilde{T}_n$ and the censoring indicators $D_i = I\{\tilde{T}_i = T_i\}$

b) We have independent censoring if an individual who is at risk at time t (i.e. $\tilde{T}_i \geq t$) has the same probability of experiencing the event in $[t, t+dt)$ as would have been the case in the situation without censoring.

A bit more formally:

$$P(t \leq \tilde{T}_i < t+dt, D_i = 1 \mid \tilde{T}_i \geq t, \text{past})$$

$$= P(t \leq T_i < t+dt \mid T_i \geq t)$$

cf ABG page 30

c) We may estimate $S(t)$ by the Kaplan-Meier estimator

$$\hat{S}(t) = \prod_{\{i : \tilde{T}_i \leq t, D_i = 1\}} \left(1 - \frac{1}{Y(\tilde{T}_i)}\right)$$

where $Y(t) = \# \{ i : \tilde{T}_i \geq t \}$ is the number at risk at time t.

The pth fractile ξ_p of the survival distribution is given by $S(\xi_p) = 1 - p$, cf page 95 in the ABG-book. The fractile is estimated by

$$\hat{\xi}_p = \inf \{ t : \hat{S}(t) \leq 1 - p \}$$

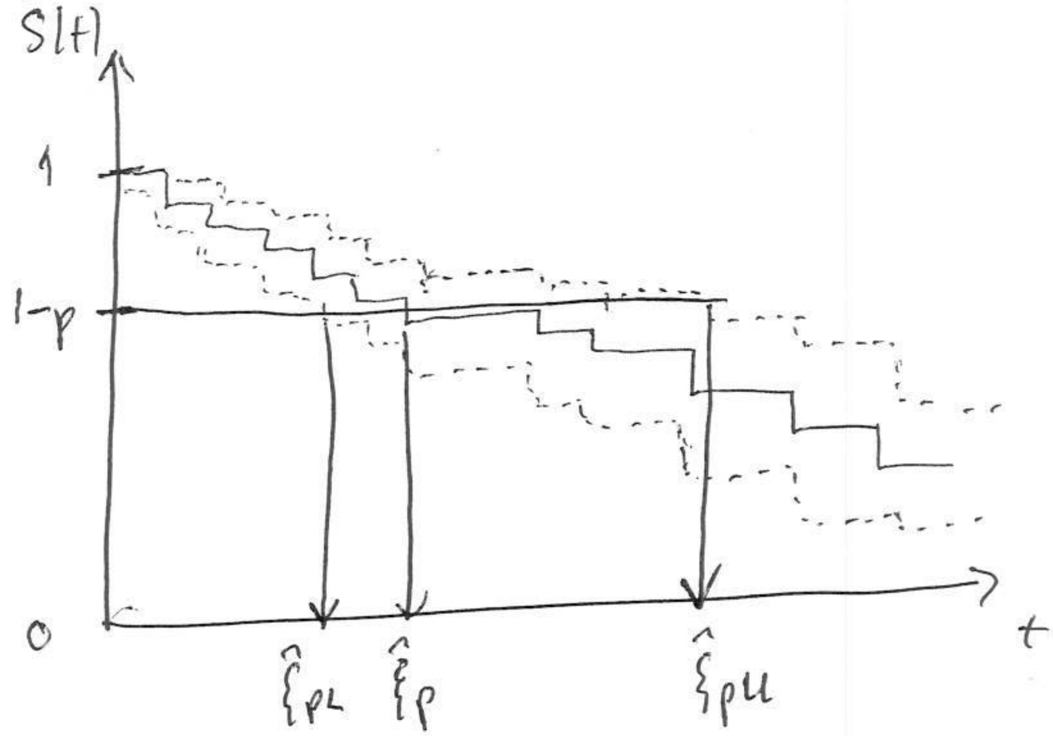
The quartiles are obtained for $p = 0.25$, $p = 0.50$ and $p = 0.75$

The lower and upper confidence limits, $\hat{\xi}_{pL}$ and $\hat{\xi}_{pU}$, are obtained in a similar manner from the confidence limits of the survival function. If we use the standard confidens limits for $S(t)$, i.e. $\hat{S}(t) \pm 1.96\,\hat{\tau}(t)$, we have

$$\hat{\xi}_{pL} = \inf\{t : \hat{S}(t) - 1.96\,\hat{\tau}(t) \leq 1-p\}$$

$$\hat{\xi}_{pU} = \inf\{t : \hat{S}(t) + 1.96\,\hat{\tau}(t) \leq 1-p\}$$

For further details, see section 3.2.3 and exercise 3.8 in the ABG-book. The sketch below illustrates the procedure



d) Output from R:

time	n.risk	n.event	survival	std.err	lower 95% CI	upper 95% CI
9	32	1	0.969	0.0308	0.798	0.996
11	31	1	0.938	0.0428	0.773	0.984
12	30	1	0.906	0.0515	0.737	0.969
20	29	2	0.844	0.0642	0.665	0.932
22	27	1	0.813	0.0690	0.629	0.911
25	26	2	0.750	0.0765	0.562	0.866
28	23	2	0.685	0.0826	0.493	0.816
31	21	1	0.652	0.0849	0.460	0.790
35	20	2	0.587	0.0880	0.397	0.736
46	18	1	0.554	0.0890	0.366	0.707
49	17	1	0.522	0.0895	0.336	0.678

We find $\hat{\xi}_{0.75} = 25$ days

$\hat{\xi}_{0.75,L} = 12$ days

$\hat{\xi}_{0.25,U} = 35$ days

Exam STK4080 December 2008 Problem 2

$N_1(t)$ and $N_2(t)$ are counting processes with intensity processes of the multiplicative form $\lambda_1(t) = \alpha_1(t) Y_1(t)$ and $\lambda_2(t) = \alpha_2(t) Y_2(t)$.

We want to test

$$H_0: \alpha_1(t) = \alpha_2(t) \quad \text{for all } t \in [0, t_0]$$

a) We will base the test on a statistic of the form

$$Z_1(t_0) = \int_0^{t_0} L(t) \left(d\hat{A}_1(t) - d\hat{A}_2(t) \right) \qquad (*)$$

where $L(t) \geq 0$ is a weight process and $\hat{A}_1(t)$ and $\hat{A}_2(t)$ are the Nelson-Aalen estimators

If H_0 is true we will have $d\hat{A}_1(t)$ and $d\hat{A}_2(t)$ fairly equal (on the average), and $Z_1(t_0)$ will be close to zero (in fact, it will have mean zero as shown in b).

If, however, $\alpha_1(t) > \alpha_2(t)$ [or $\alpha_1(t) < \alpha_2(t)$] then $d\hat{A}_1(t)$ will tend to be larger [smaller] than $d\hat{A}_2(t)$, and $Z_1(t_0)$ will be quite a bit larger [smaller] than zero.

Thus $Z_1(t_0)$ is a reasonable test statistic for "non-crossing" hazards alternatives

b) For $h = 1, 2$, the Nelson–Aalen estimators are given by

$$\hat{A}_h(t) = \int_0^t \frac{J_h(u)}{Y_h(u)} \, dN_h(u)$$

where $J_h(u) = I(Y_h(u) > 0)$.

It is assumed that the weight process $L(t) = 0$ whenever at least one of $Y_1(t)$ and $Y_2(t)$ is zero.

Then we may reformulate $Z_1(t_0)$ as [cf pp 105–106 in ABG]

$$Z_1(t_0) = \int_0^{t_0} L(t) \left\{ \frac{J_1(t)}{Y_1(t)} \, dN_1(t) - \frac{J_2(t)}{Y_2(t)} \, dN_2(t) \right\}$$

$$= \int_0^{t_0} \frac{L(t)}{Y_1(t)} \, dN_1(t) - \int_0^{t_0} \frac{L(t)}{Y_2(t)} \, dN_2(t)$$

We write $\alpha(t)$ for the common value of $\alpha_1(t)$ and $\alpha_2(t)$ under H_0.

If H_0 holds true we have the decomposition

$$dN_h(t) = \alpha(t) Y_h(t) dt + dM_h(t)$$

for $h=1,2$, where the $M_h(t)$'s are martingales.

So if H_0 holds true we have

$$Z_1(t_0) = \int_0^{t_0} \frac{L(t)}{Y_1(t)} \{\alpha(t) Y_1(t)\, dt + dM_1(t)\}$$

$$- \int_0^{t_0} \frac{L(t)}{Y_2(t)} \{\alpha(t) Y_2(t) dt + dM_2(t)\}$$

$$= \int_0^{t_0} \frac{L(t)}{Y_1(t)} dM_1(t) - \int_0^{t_0} \frac{L(t)}{Y_2(t)} dM_2(t)$$

Thus $Z_1(t_0)$, considered as a

process in t_0, is a difference of two stochastic integrals. Since a stochastic integral is itself a mean zero martingale, it follows that $Z_1(t_0)$ is a mean zero martingale.

To derive an estimator for the variance of $Z_1(t_0)$ under H_0, we consider the predictable variation process of $Z_1(t_0)$. This becomes

$$\langle Z_1 \rangle (t_0) = \int_0^{t_0} \left(\frac{L(t)}{Y_1(t)} \right)^2 \alpha(t) Y_1(t) dt$$

$$+ \int_0^{t_0} \left(\frac{L(t)}{Y_2(t)} \right)^2 \alpha(t) Y_2(t) dt$$

cf (7.48) in the ABG-book

Now $\langle Z_1 \rangle(t_0)$ simplifies to

$$\langle Z_1 \rangle(t_0) = \int_0^{t_0} \frac{L(t)^2}{Y_1(t)} \alpha(t)dt + \int_0^{t_0} \frac{L(t)^2}{Y_2(t)} \alpha(t)dt$$

$$= \int_0^{t_0} \frac{L(t)^2 \, (Y_2(t) + Y_1(t))}{Y_1(t) Y_2(t)} \alpha(t)dt$$

$$= \int_0^{t_0} \frac{L(t)^2 \, Y_\bullet(t)}{Y_1(t) Y_2(t)} \alpha(t)dt \qquad (\#)$$

where $Y_\bullet(t) = Y_1(t) + Y_2(t)$

To obtain an estimator for the variance of $Z_1(t_0)$, we replace $\alpha(t)dt$ in $(\#)$ by $d\hat{A}(t)$, where $\hat{A}(t) = \int_0^t \frac{J(u)}{Y_\bullet(u)} dN_\bullet(u)$ is the Nelson-Aalen estimator based on the aggregated

process $N_\circ(t) = N_1(t) + N_2(t)$
and $J(u) = I(Y_\circ(u) > 0)$.

We then obtain the variance estimator

$$V_{11}(t_0) = \int_0^{t_0} \frac{L(t)^2\, Y_\circ(t)}{Y_1(t)\, Y_2(t)} \frac{dN_\circ(t)}{Y_\circ(t)}$$

$$= \int_0^{t_0} \frac{L(t)^2}{Y_1(t)\, Y_2(t)}\, dN_\circ(t)$$

One may show that this estimator is unbiased under H_0 (exercise 3.10)

c) The log-rank test corresponds to the weight process
$L(t) = Y_1(t)\, Y_2(t) \,/\, Y_\circ(t)$

Then $Z_1(t_0)$ may be rewritten

as

$$Z_1(t_0) = \int_0^{t_0} \frac{L(t)}{Y_1(t)}\, dN_1(t) - \int_0^{t_0} \frac{L(t)}{Y_2(t)}\, dN_2(t)$$

$$= \int_0^{t_0} \frac{Y_2(t)}{Y_\bullet(t)}\, dN_1(t) - \int_0^{t_0} \frac{Y_1(t)}{Y_\bullet(t)}\, dN_2(t)$$

$$= \int_0^{t_0} \frac{Y_2(t)}{Y_\bullet(t)}\, dN_1(t) - \int_0^{t_0} \frac{Y_1(t)}{Y_\bullet(t)} \{dN_\bullet(t) - dN_1(t)\}$$

$$= \int_0^{t_0} \left(\frac{Y_2(t)}{Y_\bullet(t)} + \frac{Y_1(t)}{Y_\bullet(t)}\right) dN_1(t) - \int_0^{t_0} \frac{Y_1(t)}{Y_\bullet(t)}\, dN_\bullet(t)$$

$$= \int_0^{t_0} dN_1(t) - \int_0^{t_0} \frac{Y_1(t)}{Y_\bullet(t)}\, dN_\bullet(t)$$

$$= N_1(t_0) - E_1(t_0)$$

where $E(t_0) = \int_0^{t_0} \frac{Y_1(t)}{Y_\bullet(t)}\, dN_\bullet(t)$

$E_1(t_0)$ may be interpreted as the expected number of events under H_0

For the data example we get the R output

```
          N Observed Expected (O-E)^2/E (O-E)^2/V
treat=0 32        5     10.2      2.65      5.49
treat=1 32       15      9.8      2.75      5.49

 Chisq= 5.5  on 1 degrees of freedom, p= 0.0192
```

Here "treat =0" corresponds to group 1 and "treat = 1" to group 2.

We see from the output that

$N_1(t_0) = 5$ and $E_1(t_0) = 10.2$

The test statistic

$$X^2 = \frac{Z_1(t_0)^2}{V_{11}(t_0)} = \frac{(N_1(t_0) - E_1(t_0))^2}{V_{11}(t_0)}$$

takes the value 5.5, which should be compared with a chi-square distribution with 1 degree of freedom. This gives the P-value 1.9%, so the difference between the groups is significant

Since $N_1(t_0) = 5$ is smaller than $E_1(t_0) = 10.2$, the treatment with both MTX and CSP reduce the risk of the life threatening complication.

Exam STK4080 December 2008 Problem 3

We have counting processes $N_1(t), \ldots, N_n(t)$ with intensity processes

$$\lambda_i(t) = Y_i(t)\{\beta_0(t) + \beta_1(t)x_{i1} + \beta_2(t)x_{i2}\}$$

where $Y_i(t)$ is an at risk indicator $(0,1)$

a) We introduce $B(t) = (B_0(t), B_1(t), B_2(t))^T$, where $B_j(t) = \int_0^t \beta_j(u)du$. We also introduce the vectors

$$N(t) = (N_1(t), \ldots, N_n(t))^T$$

and

$$M(t) = (M_1(t), \ldots, M_n(t))^T,$$

where $M_i(t) = N_i(t) - \int_0^t \lambda_i(u)du,$

and the matrix

$$X(t) = \begin{pmatrix} Y_1(t) & Y_1(t)\,x_{11} & Y_2(t)\,x_{12} \\ Y_2(t) & Y_2(t)\,x_{21} & Y_2(t)\,x_{22} \\ \vdots & \vdots & \vdots \\ Y_n(t) & Y_n(t)\,x_{n1} & Y_n(t)\,x_{n2} \end{pmatrix}$$

For each $i = 1, \dots, n$ we have the decomposition

$$dN_i(t) = \lambda_i(t)\,dt + dM_i(t)$$

$$= Y_i(t)\left\{\beta_0(t) + \beta_1(t)\,x_{i1} + \beta_2(t)\,x_{i2}\right\}dt + dM_i(t)$$

$$= Y_i(t)\,dB_0(t) + Y_i(t)\,x_{i1}\,dB_1(t) + Y_i(t)\,x_{i2}\,dB_2(t) + dM_i(t)$$

These n relations may be written in vector / matrix form as follows

$$dN(t) = X(t)\, dB(t) + dM(t) \quad (1)$$

Equation (1) has the form of a linear regression model with $dM(t)$ as the noise term

We may use ordinary least squares to find the estimator for $dB(t)$, namely

$$d\hat{B}(t) = (X(t)^T X(t))^{-1} X(t)^T dN(t), \quad (2)$$

provided that $X(t)$ has full rank

It is convenient to introduce

the least squares generealized inverse

$$X^{-}(t) = (X(t)^{T} X(t))^{-1} X(t)^{T} \qquad (3)$$

and the indicator

$$J(t) = I\{X(t) \text{ has full rank}\}$$

Then we obtain the estimator for $B(t)$ by accumulating the increments (2) over all time-points when $X(t)$ has full rank. By using the notation (3), we then obtain

$$\hat{B}(t) = \int_0^t J(u)\, X^{-}(u)\, dN(u) \qquad (4)$$

If we let $T_1 < T_2 < \ldots$ be the

times when events occur (i.e. the jump times of the $N_i(t)$'s), we may write

$$\hat{B}(t) = \sum_{T_j \le t} J(T_j)\, X^-(T_j)\, \Delta N(T_j)$$

To see that $\hat{B}(t)$ is approximately unbiased, we insert (9) in (4) to obtain

$$\hat{B}(t) = \int_0^t J(u)\, X^-(u) \{ X(u)\, dB(u) + dM(u) \}$$

$$= \int_0^t J(u)\, X^-(u)\, X(u)\, dB(u)$$

$$+ \int_0^t J(u)\, X^-(u)\, dM(u)$$

By (13) we see that

$X^{-}(u)\, X(u) = I$, the identity matrix. Hence we have that

$$\hat{B}(t) = \int_0^t J(u)\, dB(u) + \int_0^t J(u) X^{-}(u)\, dM(u)$$

Here the last integral is a (vector valued) stochastic integral, and hence has mean zero.

Therefore

$$E\,\hat{B}(t) = E\left\{ \int_0^t J(u)\, dB(u) \right\}$$

$$= \int_0^t E(J(u))\, dB(u) = \int_0^t P(J(u) > 0)\, dB(u)$$

$$\approx \int_0^t dB(u) = B(t)$$

So $\vec{B}(t)$ is approximately unbiased when

$$P(J(u)>0) = P(X(u) \text{ has full rank})$$

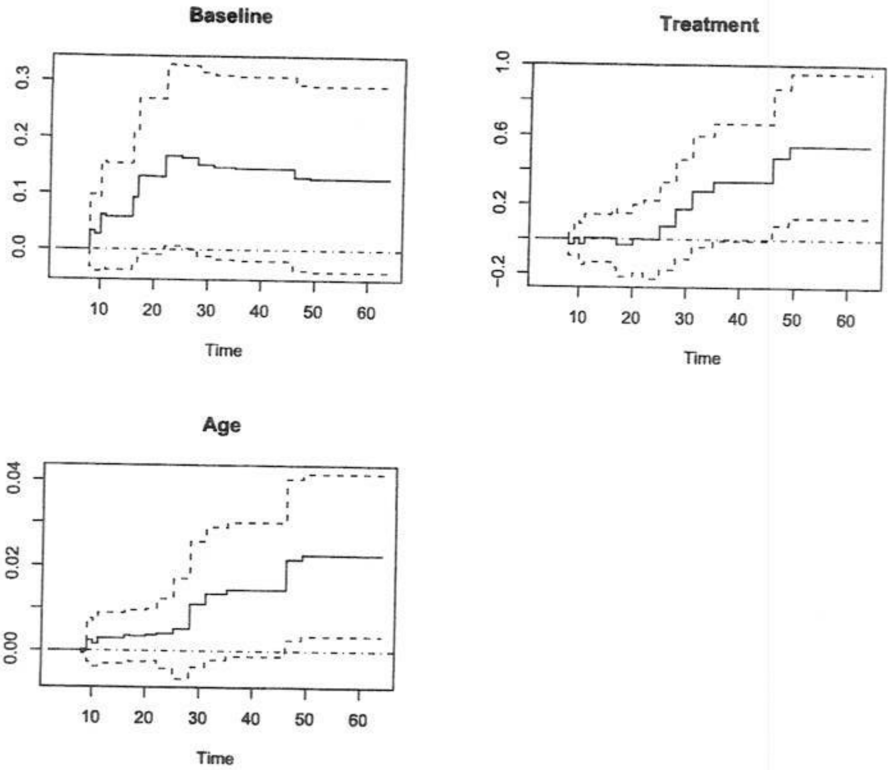
is close to 1 for $0<u<t$.

We now consider the study described in Problem 1 and introduce the covariates

$$x_{i1} = \begin{cases} 0 & \text{if both MTX and CSP} \\ 1 & \text{if only MTX} \end{cases}$$

$$x_{i2} = \text{age} - 20$$

A fit of the model gives the plots on the next page



The cumulative baseline estimate is for a patient aged 20 years who is treated with both MTX and CSP. We see that such a patient has a fairly high risk of experiencing the complication the first 20 days after transplantation. The hazard is about $0.15/20 = 0.008$ per day the first 20 days (estimated roughly from the slope of the estimate of the cumulative hazard). After 20-25 days the risk of the complication is low (since the slope is about zero).

The cumulative effect of treatment is about zero the first 20 days, so over this period, there is no difference between the treatments. But after

20-25 days the estimate for treatment increases with a slope of about $0.50/(50-25) = 0.02$. This means that after 20-25 days the risk of the complication is substantially larger for patients who get only MTX.

Finally the estimate for age shows that the risk of the complication increases with age. A rough estimate is that the hazard increases with $0.02/50 = 0.0004$ per year, or 0.004 per 10 years

Exam STK 4080 December 2008 Problem 4

$M = \{M_0, M_1, \dots\}$ is a martingale relative to the history $\{\mathcal{F}_n\}$ and $M_0 = 0$.

a) That M is a martingale means that it satisfies the martingale property

$$E(M_n \mid \mathcal{F}_m) = M_m$$

for all $m < n$. (It is sufficient that it holds for $m = n-1$.)
By using the rule of double expectation we obtain ($m < n$)

$$E(M_n) = E\{E(M_n \mid \mathcal{F}_m)\} = E(M_m)$$

In particular this gives for all $n > 0$

$$E(M_n) = E(M_0) = 0$$

b) The predictable variation process of M is given by $\langle M\rangle_0 = 0$ and

$$\langle M\rangle_n = \sum_{i=1}^{n} \operatorname{Var}(\Delta M_i \mid \mathcal{F}_{i-1})$$

where $\Delta M_i = M_i - M_{i-1}$ is the martingale difference. Note that since $E(\Delta M_i \mid \mathcal{F}_{i-1}) = 0$ (by the martingale property), we may also write

$$\langle M\rangle_n = \sum_{i=1}^{n} E((\Delta M_i)^2 \mid \mathcal{F}_{i-1})$$

To show that $M_n^2 - \langle M\rangle_n$ is a martingale, it suffices to show that

$$E(M_n^2 - \langle M \rangle_n \mid \mathcal{F}_{n-1}) = M_{n-1}^2 - \langle M \rangle_{n-1}$$

Now

$$M_n^2 - \langle M \rangle_n = (M_{n-1} + \Delta M_n)^2$$
$$- \langle M \rangle_{n-1} - \Delta \langle M \rangle_n$$

Where

$$\Delta \langle M \rangle_n = \langle M \rangle_n - \langle M \rangle_{n-1}$$
$$= E((\Delta M_n)^2 \mid \mathcal{F}_{n-1})$$

Hence we have

$$E(M_n^2 - \langle M \rangle_n \mid \mathcal{F}_{n-1})$$

$$= E(M_{n-1}^2 + 2 M_{n-1} \Delta M_n + (\Delta M_n)^2$$
$$- \langle M \rangle_{n-1} - \Delta \langle M \rangle_n \mid \mathcal{F}_{n-1})$$

$$= M_{n-1}^2 + 2 M_{n-1} \underbrace{E(\Delta M_n \mid \mathcal{F}_{n-1})}_{=0}$$

$$+ E((\Delta M_n)^2 \mid \mathcal{F}_{n-1})$$

$$- \langle M \rangle_{n-1} - E((\Delta M_n)^2 \mid \mathcal{F}_{n-1})$$

$$= M_{n-1}^2 - \langle M \rangle_{n-1}$$

This shows that $M_n^2 - \langle M \rangle_n$ is a mean zero martingale:

Now, since $EM_n = 0$, we have

$$\mathrm{Var}(M_n) = E(M_n^2)$$

$$= E(M_n^2 - \langle M \rangle_n + \langle M \rangle_n)$$

$$= E(M_n^2 - \langle M \rangle_n) + E\langle M \rangle_n$$

$$= 0 + E\langle M \rangle_n = E\langle M \rangle_n$$

c) We let $H = \{H_0, H_1, H_2, \dots\}$ be predictable process, which means that H_n is known at time $n-1$.

The transformation $Z = H \bullet M$ is given by $Z_0 = H_0 M_0 = 0$ and

$$Z_n = \sum_{i=1}^{n} H_i \,\Delta M_i$$

$$= H_1 (M_1 - M_0) + H_2 \cdot (M_2 - M_1) + \dots + H_n (M_n - M_{n-1})$$

Now we have

$$E(Z_n | \mathcal{F}_{n-1}) = E(Z_{n-1} + H_n \Delta M_n \mid \mathcal{F}_{n-1})$$

$$= Z_{n-1} + E(H_n \Delta M_n \mid \mathcal{F}_{n-1})$$

$$= Z_{n-1} + H_n \underbrace{E(\Delta M_n \mid \mathcal{F}_{n-1})}_{=0} = Z_{n-1}$$

So Z_n is a mean zero martingale.

The predictable variation process of Z is given by $\langle Z \rangle_0 = 0$ and

$$\langle Z \rangle_n = \sum_{i=1}^{n} E\{(\Delta Z_i)^2 \mid \mathcal{F}_{i-1}\}$$

Now we have

$$\Delta Z_i = Z_i - Z_{i-1} = H_i \Delta M_i$$

It follows that for $n \geq 1$ we have

$$\begin{aligned}
\langle Z \rangle_n &= \sum_{i=1}^{n} E\{(H_i \Delta M_i)^2 \mid \mathcal{F}_{i-1}\} \\
&= \sum_{i=1}^{n} E\{H_i^2 (\Delta M_i)^2 \mid \mathcal{F}_{i-1}\} \\
&= \sum_{i=1}^{n} H_i^2 E\{(\Delta M_i)^2 \mid \mathcal{F}_{i-1}\} \\
&= \sum_{i=1}^{n} H_i^2 \Delta \langle M \rangle_i
\end{aligned}$$

where the second last equality follows since H is predictable