M3F22/M4F22/M5F22 EXAMINATION SOLUTIONS 2017-18

Q1 (*Lack of memory and the exponential laws*).
(a) Consider a probability distribution (law) F on $(0, \infty)$, interpreted as the lifetime law of components, say. Then F has the *lack-of-memory property* iff the components show no aging – that is, if a component still in use behaves as if new. The condition for this is (with X the age of the current component)

$$P(X > s + t | X > s) = P(X > t) \qquad (s, t > 0):$$

$$P(X > s + t) = P(X > s)P(X > t).$$

Writing $\overline{F}(x) := 1 - F(x)$ $(x \geq 0)$ for the *tail* of F, this says that

$$\overline{F}(s + t) = \overline{F}(s)\overline{F}(t) \qquad (s, t \geq 0). \qquad \textbf{[5]}$$

(b) Obvious solutions are

$$\overline{F}(t) = e^{-\lambda t}, \qquad F(t) = 1 - e^{-\lambda t}$$

for some $\lambda > 0$ – the exponential law $E(\lambda)$. Now

$$f(s + t) = f(s)f(t) \qquad (s, t \geq 0) \qquad (CFE)$$

is a 'functional equation' – the *Cauchy functional equation* (CFE) – and (we quote) these are the *only* bounded solutions, (indeed, the only ones subject to any – minimal – regularity condition).

So the exponential laws $E(\lambda)$ are *characterized* by the lack-of-memory property. **[5]**
(c) The Poisson point process $Ppp(\lambda)$ with rate $\lambda > 0$ is defined to have the inter-arrival times independent $E(\lambda)$. It is the lack-of-memory property of the $E(\lambda)$ that makes the Poisson process the basic model for events occurring 'out of the blue'. Typical examples are accidents, insurance claims, hospital admissions, earthquakes, volcanic eruptions etc. **[5]**
(d) *Limitations*. The weakness in this model for insurance claims is that a major catastrophe produces a *cluster* of claims. The independence assumption will fail badly *within* clusters, though it may still work well *between* clusters. **[5]**
[(a)-(c): seen – lectures; (d): mainly unseen]

Q2 (*No-Arbitrage Theorem (NA Theorem)*).
(a) *Proof*: $\Leftarrow$. In discrete time: we take the state space Ω to be discrete also; we can then retain only sample points ω with positive probability, $P(\omega) > 0$.

Assume such an equivalent martingale measure (EMM) P^* exists. For any self-financing strategy H, we have

$$\tilde{V}_n(H) = V_0(H) + \Sigma_1^n H_j.\Delta\tilde{S}_j$$

(at the jth trade, the gain in value $\Delta V_j(H)$ is the amount H_j of the jth asset that we buy, times the gain ΔS_j in its price; similarly for $\tilde{V}_j$, $\tilde{S}_j$ with discounting). This gives $\tilde{V}_n(H)$ as the *martingale transform* of the P^*-martingale $\tilde{S}_j$ by $H = (H_n)$, so $\tilde{V}_n(H)$ is a P^*-martingale. So the initial and final P^*-expectations are the same: using E^* for P^*-expectation,

$$E^*[\tilde{V}_N(H)] = E^*[\tilde{V}_0(H)].$$

If the strategy is admissible and its initial value – the RHS above – is zero, the LHS $E^*[\tilde{V}_N(H)]$ is zero, but $\tilde{V}_N(H) \geq 0$ (by admissibility). Since each $P(\{\omega\}) > 0$ (by assumption), each $P^*(\{\omega\}) > 0$ (by equivalence). This and $\tilde{V}_N(H) \geq 0$ force $\tilde{V}_N(H) = 0$ (sum of non-negatives can only be 0 if each term is 0). So no arbitrage is possible. // **[6]**

(b) The direct half (no arbitrage implies existence of an EMM) needs the *Separating Hyperplane Theorem*. The general form of this is related to the Hahn-Banach Theorem of Functional Analysis, which needs the Axiom of Choice (AC). In a finite-dimensional setting (as in (i)), one can use Euclidean geometry – much simpler. **[3]**

(c) The NA Theorem (NA iff EMMs *exist*) shows that the assumption of NA is needed to be able to *price* assets, including *options*. (Completeness is needed to make EMMs, and so prices, *unique*; real markets are incomplete; real prices are non-unique; "You'd better shop around".) In particular, one can price options without needing to know the market participant's *utility* function – i.e., his attitude to risk. This is the *Arbitrage Pricing Technique (APT)*, due to the late Steve (S. A.) Ross (1976/78): it takes the qualitative insight of the NA Theorem above, and uses it systematically to produce *quantitative* results – asset pricing, etc. (EMMs correspond to *pricing kernels*). **[3]**

(d) Arbitrage opportunities do exist in reality – and professional arbitrageurs hunt for them. They are a 'second-order effect': anyone opening himself to arbitrage is in effect offering the market free money (being used as a 'money-pump'); the market will take the free money without limit until he withdraws from the market (bankrupt or otherwise), or at least withdraws the arbitrage opportunity – which is thus 'arbitraged away'. **[4]**

(e) With EMMs, we can price assets (albeit non-uniquely without completeness – to within an interval, the 'bid-ask spread'). But without NA and EMMs, pricing cannot be done systematically at all. If assets cannot be *priced* reliably, they will not be *traded*, in any significant quantity. So option exchanges (such as CBOE), where options can be traded in quantity and so as liquid assets, could not have been developed. So the existence of a mass market in options and other assets (an essential aspect of the City of London and other global financial centres) depends on the no-arbitrage assumption.

With 'money-pumps' as in (d) available on a large scale, one would have a *disorderly market* (the economy would be like a bath with taps running but no plug). **[4]**

[(a)-(d): seen; (e): unseen]

Q3 (*Theta*). Given

$$Ke^{-r(T-t)}\phi(d_2) = S\phi(d_1): \qquad (*)$$

(a) *Calls.* Given the Black-Scholes formula for the price c_t of European calls,

$$c_t = S_t\Phi(d_1) - Ke^{-r(T-t)}\Phi(d_2),$$

$$d_{1,2} := [\log(S/K) + (r \pm \frac{1}{2}\sigma^2)(T-t)]/\sigma\sqrt{T-t}: \qquad d_2 = d_1 - \sigma\sqrt{T-t}:$$

(i) Differentiating and using $(*)$: as

$$\partial(d_1 - d_2)/\partial t = \partial(\sigma\sqrt{T-t})/\partial t = -\frac{1}{2}\sigma/\sqrt{T-t}:$$

$$\Theta = \partial c_t/\partial t = S\phi(d_1)\frac{\partial d_1}{\partial t} - rKe^{-r(T-t)}\Phi(d_2) - Ke^{-r(T-t)}\phi(d_2)\frac{\partial d_2}{\partial t}:$$

$$\Theta = Ke^{-r(T-t)}[\phi(d_2)\frac{\partial(d_1-d_2)}{\partial t} - r\Phi(d_2)]:$$

$$\Theta = -Ke^{-r(T-t)}[\phi(d_2)\cdot\frac{\frac{1}{2}\sigma}{\sqrt{T-t}} + r\Phi(d_2)] < 0. \qquad \textbf{[6]}$$

(ii) Interpretation: an option is (partly) an insurance against future uncertainty. As time passes, there is less future (till expiry) to protect against, so such protection becomes less valuable. **[4]**

(b) *Puts.* Given the corresponding BS formula for European puts,

$$p_t = Ke^{-r(T-t)}\Phi(-d_2) - S_t\Phi(-d_1),$$

(i) As above, as $\phi(-x) = \phi(x)$,

$$\Theta = \partial p_t/\partial t = rKe^{-r(T-t)}\Phi(-d_2) + Ke^{-r(T-t)}\phi(d_2)\frac{\partial(-d_2)}{\partial t} - S\phi(d_1)\frac{\partial(-d_1)}{\partial t}:$$

$$\Theta = Ke^{-r(T-t)}[r\Phi(-d_2)+\phi(d_2)\frac{\partial(d_1-d_2)}{\partial t}] = Ke^{-r(T-t)}[r\Phi(-d_2)-\phi(d_2)\cdot\frac{\frac{1}{2}\sigma}{\sqrt{T-t}}].$$

This can change sign! **[6]**

Or: use *put-call parity*: $S + P - C = Ke^{-r(T-t)} = Ke^{-rT}e^{rt}$. So $\Theta_P = \Theta_C + rKe^{-rT}e^{rt}$. The first term is < 0 by (a), the second is > 0, so the sum can change sign.

(ii) The situation with puts is *different*, because of the different role of the strike K (fixed, while S varies).

For large enough K (when a put option – the right to *sell* at price K – will be deeply in the money), the option stands to make a large profit. So the more time passes, the nearer this is to being realised, the better, so $\Theta > 0$.

This is the situation when $K >> S$, S/K small (positive), $\log(S/K)$, d_1, d_2 small (near $-\infty$), $\Phi(-d_2)$ near 1, but $\phi(-d_2)$ exponentially small, so negligible. So the second (negative) term is negligible, and the first (positive) term predominates. **[4]**

Note. We have used S_t for the stock price at time t in the Black-Scholes formulae, but abbreviated this to S in our working. There is no need for a "$\partial S/\partial t$" term in the calculus! Indeed, there can't be one (the SDE for GBM in Q5 involves Brownian motion, and this is not differentiable). There is a functional dependence on time t in the discounting multiplying K, and in d_1, d_2. There is no functional dependence of S on t (although of course the stock price *varies* with time). In mathematical terms, what S depends on is the *randomness*, ω (suppressed as usual in our notation), as S is a *stochastic process* (random function of time). In financial terms, what S depends on is the *market*.

[(ai), (bi): similar seen (for vega, the derivative wrt the volatility σ, and ρ, the derivative wrt the riskless interest-rate r); (aii), (bii): unseen]

Q4 (*Renewal theory and ruin theory*).
(a) *Safety loading.* With $c > 0$ the premium rate at which cash comes in, $\lambda > 0$ the rate at which claims occur, $\mu \in (0, \infty)$ the mean claim size, cash goes out at rate $\lambda\mu$, so one needs ('more in than out') $c > \lambda\mu$. The *safety loading* $\rho > 0$ is defined by

$$\frac{c}{\lambda\mu} = 1 + \rho. \qquad (SL) \quad [\mathbf{3}]$$

(b) *Key renewal theorem.* The *renewal equation* for F and z (both known) is the *integral equation*

$$Z(t) = z(t) + \int_0^t Z(t-u)dF(u) \quad (t \geq 0): \quad Z = z + F * Z. \qquad (RE)$$

Here F (the lifetime distribution) and z are given, and (RE) is to be solved for Z. Then for $U := \sum_0^\infty F^{*n}$ the renewal function of F:

Theorem (Key Renewal Theorem; W. L. Smith). If z in (RE) is directly Riemann integrable, then with U the renewal function of F,

$$\lim_{t\to\infty} Z(t) = \lim_{t\to\infty}(U * z)(t) = \frac{1}{\mu}\int_0^\infty z(x)dx. \qquad [\mathbf{3}]$$

(c) *The Lundberg (or adjustment) coefficient*, r. This is the point $r > 0$ (assumed to exist – a strengthening of the Small Claims Condition; it is then unique) such that the MGF of $Z = Z_1$ satisfies, writing M for M_{X_1} for short,

$$M_{Z_1}(r) := E[\exp\{r(X_1 - cW_1)\}] = M(r).\frac{\lambda}{\lambda + cr} = 1: \quad M(r) = 1 + \frac{cr}{\lambda}$$

(the product by independence, the second factor as $W_1 \sim E(\lambda)$).

The bigger r is, the better. For (from the graph of M): the bigger r is, the bigger the strip of holomorphy of the claim-size MGF, so the smaller the claim-size tails, so the smaller the chance of a damaging big claim. **[4]**

(d) *The Esscher transform.* By above,

$$M(r) := \int_0^\infty e^{rx}dF(x) = -\int_0^\infty e^{rx}d(1-F)(x) = 1 + \frac{cr}{\lambda}.$$

Integrating by parts, the integrated term is 1, giving

$$\int_0^\infty (1 - F(x))e^{rx}dx = \frac{c}{\lambda}, = (1+\rho)\mu,$$

by (SL). So

$$\frac{\lambda}{c}(1-F(x))e^{rx} = \frac{1}{(1+\rho)\mu}(1-F(x))e^{rx}$$

is a probability density on $(0,\infty)$ – of G, say. Then $F \mapsto G$ is called the *Esscher transform.* **[3]**

(e) *The Cramér estimate of ruin.* Given the integral equation for the ruin probability $\psi(u)$:

$$\psi(u) = \frac{1}{(1+\rho)}\int_u^\infty \frac{(1-F(x))}{\mu}dx + \frac{1}{(1+\rho)}\cdot\int_0^u \psi(u-x)\frac{(1-F(x))}{\mu}dx \quad (*)$$

(as $(1-F(x))/\mu$ is a probability density, so integrates to 1). This is of renewal-equation type, *except* that, as $(1-F(x))/\mu$ is a probability *density*, the factor $1/(1+\rho)<1$ turns it into a *sub-probability* (or *defective*) density.

Theorem (Cramér's estimate of ruin, 1930).

For the Cramér-Lundberg model, with Lundberg coefficient $r>0$ and $\psi(u)$ the probability of ruin with initial capital u,

$$e^{ru}\psi(u)\to C: \qquad \psi(u)\sim Ce^{-ru} \qquad (u\to\infty),$$

with C an (identifiable) constant. That is, as the initial capital *increases*, the ruin probability *decreases exponentially.*

Proof. Multiply $(*)$ by e^{ru}, and regard it as an integral equation in $\psi(u)e^{ru}$:

$$[\psi(u)e^{ru}] = e^{ru}\int_u^\infty \frac{(1-F(x))}{(1+\rho)\mu}dx + \int_0^u [\psi(u-x)e^{r(u-x)}]\frac{e^{rx}(1-F(x))}{(1+\rho)\mu}dx.$$

This is now an integral equation of renewal type (RE). So by the Key Renewal Theorem, its solution $\psi(u)e^{ru}$ has a limit, C say, as $u\to\infty$ (C can be read off from the Key Renewal Theorem). // **[4]**

(f) The most unrealistic assumption here is that the claims are independent. A natural disaster will produce a cluster of claims, heavily dependent. This can be handled by treating the clusters as 'points' in a Poisson process. **[3]**
[(a) - (e): Seen – lectures; (f): unseen]

Q5 (Mastery question: *Geometric Brownian motion and its quadratic variation*).

(a) Consider the process

$$X_t = f(t, B_t) := x_0. \exp\{(\mu - \frac{1}{2}\sigma^2)t + \sigma B_t\} : \ \log X_t = const + (\mu - \frac{1}{2}\sigma^2)t + \sigma B_t, \qquad (*)$$

with $B = (B_t)$ Brownian motion (BM). Here, since

$$f(t, x) = x_0. \exp\{(\mu - \frac{1}{2}s^2)t + \sigma x\},$$

$$f_1 = (\mu - \frac{1}{2}\sigma^2)f, \qquad f_2 = \sigma f, \qquad f_{22} = \sigma^2 f.$$

By Itô's Lemma,

$$dX_t = f_1 dx + f_2 dB_t + \frac{1}{2} f_{22}(dB_t)^2 :$$

$$dX_t = df = [(\mu - \frac{1}{2}\sigma^2)f + \frac{1}{2}\sigma^2 f]dt + \sigma f dB_t :$$

$$dX_t = \mu f dt + \sigma f dB_t = \mu X_t dt + \sigma X_t dB_t :$$

X satisfies the SDE

$$dX_t = X_t(\mu dt + \sigma dB_t) : \qquad dX_t/X_t = \mu dt + \sigma dB_t, \qquad (GBM)$$

geometric Brownian motion (GBM). It is used to model (stock) *price processes* in the Black-Scholes model – where, by $(*)$, *log-prices* $\log X_t$ *are normally distributed*, so *prices* are *log-normally distributed.* **[8]**

(b) *Interpretation.* The μdt term on the RHS corresponds to a *riskless* asset with *return* rate μ. The σdB_t term corresponds to a *risky* asset with *volatility* σ; the Brownian motion (B_t) models the uncertainty driving the economic/financial environment; the volatility σ represents how sensitive this particular stock is to this. **[3]**

(c) *Quadratic variation.* Recall $(dB_t)^2 = dt$ (Itô: differential form of Lévy's theorem on quadratic variation of BM). So

$$(dX_t)^2 = X_t^2(\mu^2(dt)^2 + 2\mu\sigma dt dB_t + \sigma^2(dB_t)^2) : \qquad (dX_t)^2 = \sigma^2 X_t^2 dt,$$

as above. So, as with BM, GBM has quadratic variation (QV)

$$\sigma^2 \int_0^t X_s^2 ds. \qquad \mathbf{[3]}$$

(d) *Expected QV*. By Fubini's theorem, it has expected QV

$$E[\sigma^2 \int_0^t X_s^2 ds] = \sigma^2 \int_0^t E[X_s^2] ds,$$

By $(*)$, as $B_t \sim \sqrt{t}Z$ with $Z \sim N(0,1)$ with MGF $\exp\{\frac{1}{2}t^2\}$,

$$X_t^2 = x_0^2 \cdot \exp\{(2\mu - \sigma^2)t\} \cdot \exp\{2\sigma B_t\} \sim x_0^2 \cdot \exp\{(2\mu - \sigma^2)t\} \cdot \exp\{2\sigma\sqrt{t}Z\}.$$

By the normal MGF, the last term has expectation $\exp\{\frac{1}{2}(2\sigma\sqrt{t})^2\} = \exp\{2\sigma^2 t\}$. Combining,

$$E[X_t^2] = x_0^2 \cdot \exp\{(2\mu - \sigma^2)t\} \cdot \exp\{2\sigma^2 t\} = x_0^2 \cdot \exp\{(2\mu + \sigma^2)t\}.$$

So the expected QV of GBM is

$$x_0^2 \sigma^2 \frac{\exp\{(2\mu + \sigma^2)t\} - 1}{2\mu + \sigma^2}. \qquad [\mathbf{6}]$$

[(a), (b): seen, lectures; (c), (d): unseen] N. H. Bingham