Society of Actuaries
Course 4 Exam Solutions
Fall 2003

Question #1
Key: A

The Yule-Walker equations are:
$.5 = \phi_1 + .5\phi_2$
$.1 = .5\phi_1 + \phi_2$
The solution is $\phi_2 = -.2$.

Question #2
Key: E

The 40th percentile is the .4(12) = 4.8th smallest observation. By interpolation it is .2(86) +.8(90) = 89.2. The 80th percentile is the .8(12) = 9.6th smallest observation. By interpolation it is .4(200) +.6(210) = 206.

The equations to solve are
$.4 = \dfrac{(89.2/\theta)^\gamma}{1+(89.2/\theta)^\gamma}$ and $.8 = \dfrac{(206/\theta)^\gamma}{1+(206/\theta)^\gamma}$.
Solving each for the parenthetical expression gives $\dfrac{2}{3} = (89.2/\theta)^\gamma$ and $4 = (206/\theta)^\gamma$.
Taking the ratio of the second equation to the first gives $6 = (206/89.2)^\gamma$ which leads to $\gamma = \ln(6)/\ln(206/89.2) = 2.1407$. Then $4^{1/2.1407} = 206/\theta$ for $\theta = 107.8$.

Question #3
Key: E

The standard for full credibility is $\left(\dfrac{1.645}{.02}\right)^2\left(1+\dfrac{Var(X)}{E(X)^2}\right)$ where X is the claim size variable.

For the Pareto variable, $E(X) = .5/5 = .1$ and $Var(X) = \dfrac{2(.5)^2}{5(4)} - (.1)^2 = .015$. Then the standard is $\left(\dfrac{1.645}{.02}\right)^2\left(1+\dfrac{.015}{.1^2}\right) = 16{,}913$ claims.

Question #4
Key: B

The kernel is a triangle with a base of 4 and a height at the middle of 0.5 (so the area is 1). The length of the base is twice the bandwidth. Any observation within 2 of 2.5 will contribute to the estimate. For the observation at 2, when the triangle is centered at 2, the height of the triangle at 2.5 is .375 (it is one-quarter the way from 2 to the end of the triangle at 4 and so the height is one-quarter the way from 0.5 to 0). Similarly the points at 3 are also 0.5 away and so the height of the associated triangle is also .375. Each triangle height is weighted by the empirical probability at the associated point. So the estimate at 2.5 is (1/5)(3/8) + (3/5)(3/8) + (1/5)(0) = 12/40.

Question #5
Key: D

The standard error is $\sqrt{92/8}$. With 4 of the *X*s equal to 1 and 6 equal to 0 the average is 0.4. Then

$$s_{\hat{\beta}}^2 = \frac{s^2}{\sum (X_i - \bar{X})^2} = \frac{92/8}{4(1-.4)^2 + 6(0-.4)^2} = 4.7917 \text{ and } t = \frac{\hat{\beta}}{s_{\hat{\beta}}} = \frac{4}{\sqrt{4.7917}} = 1.83.$$

Question #6
Key: A

The distribution function is $F(x) = \int_1^x \alpha t^{-\alpha-1} dt = -t^{-\alpha}\Big|_1^x = 1 - x^{-\alpha}$. The likelihood function is

$$\begin{aligned} L &= f(3) f(6) f(14)[1-F(25)]^2 \\ &= \alpha 3^{-\alpha-1} \alpha 6^{-\alpha-1} \alpha 14^{-\alpha-1} (25^{-\alpha})^2 \\ &\propto \alpha^3 [3(6)(14)(625)]^{-\alpha}. \end{aligned}$$

Taking logs, differentiating, setting equal to zero, and solving:

$\ln L = 3\ln\alpha - \alpha \ln 157{,}500$ plus a constant

$(\ln L)' = 3\alpha^{-1} - \ln 157{,}500 = 0$

$\hat{\alpha} = 3/\ln 157{,}500 = .2507.$

Question #7
Key: C

$$\pi(q|1,1) \propto p(1|q)p(1|q)\pi(q) = 2q(1-q)2q(1-q)4q^3 \propto q^5(1-q)^2$$

$$\int_0^1 q^5(1-q)^2\,dq = 1/168, \quad \pi(q|1,1) = 168q^5(1-q)^2.$$

The expected number of claims in a year is $E(X\,|\,q) = 2q$ and so the Bayesian estimate is

$$E(2q|1,1) = \int_0^1 2q(168)q^5(1-q)^2\,dq = 4/3.$$

The answer can be obtained without integrals by recognizing that the posterior distribution of q is beta with a = 6 and b = 3. The posterior mean is $E(q|1,1) = a/(a+b) = 6/9 = 2/3.$ The posterior mean of $2q$ is then 4/3.

Question #8
Key: D

For the method of moments estimate,

$$386 = e^{\mu+.5\sigma^2}, \quad 457,480.2 = e^{2\mu+2\sigma^2}$$

$$5.9558 = \mu + .5\sigma^2, \quad 13.0335 = 2\mu + 2\sigma^2$$

$$\hat{\mu} = 5.3949, \quad \hat{\sigma}^2 = 1.1218.$$

Then

$$\begin{aligned}
E(X \wedge 500) &= e^{5.3949+.5(1.1218)}\Phi\left(\frac{\ln 500 - 5.3949 - 1.1218}{\sqrt{1.1218}}\right) + 500\left[1-\Phi\left(\frac{\ln 500 - 5.3949}{\sqrt{1.1218}}\right)\right] \\
&= 386\Phi(-.2853) + 500[1-\Phi(.7739)] \\
&= 386(.3877) + 500(.2195) = 259.
\end{aligned}$$

Note-these calculations use exact normal probabilities. Rounding and using the normal table that accompanies the exam will produce a different numerical answer but the same letter answer.

Question #9
Key: E

Summing over i and j the least-squares quantity to minimize is

$$\sum_{i=1}^{2}\sum_{j=1}^{n}(Y_{2ij} - \delta - \phi Y_{1ij} - \theta X_{ij})^2 = \sum_{j=1}^{n}(Y_{21j} - \delta - \phi Y_{11j})^2 + \sum_{j=1}^{n}(Y_{22j} - \delta - \phi Y_{12j} - \theta)^2$$

where the sum is split for i=1 and i=2 where the X values are known. Differentiating with respect to the two variables gives the equations

$$-2(30n - \delta n - .75(40)n) - 2(37n - \delta n - .75(41)n - \theta n) = 0$$

$$-2(37n - \delta n - .75(41)n - \theta n) = 0.$$

Substituting the second equation into the first implies that $\delta = 0$. Then the second equation yields $\theta = 37 - .75(41) = 6.25.$

Question #10
Key: D

Because the values are already ranked, the test statistic is immediately calculated as the sum of the given values for Sample I: $R = 1 + 2 + 3 + 4 + 7 + 9 + 13 + 19 + 20 = 78$. The other needed values are $n = 9$ and $m = 11$, the two sample sizes. The mean is $n(n+m+1)/2 = 94.5$ and the variance is $nm(n+m+1)/12 = 173.25$. The test statistic is $Z = (78 - 94.5)/\sqrt{173.25} = -1.25$. The p-value is twice (because it is a two-tailed test) the probability of being more extreme than the test statistic, $p = 2\Pr(Z < -1.25) = .210$.

Question #11
Key: C

Let N be the Poisson claim count variable, let X be the claim size variable, and let S be the aggregate loss variable.

$$\mu(\theta) = E(S \mid \theta) = E(N \mid \theta)E(X \mid \theta) = \theta 10\theta = 10\theta^2$$

$$v(\theta) = Var(S \mid \theta) = E(N \mid \theta)E(X^2 \mid \theta) = \theta 200\theta^2 = 200\theta^3$$

$$\mu = E(10\theta^2) = \int_1^\infty 10\theta^2(5\theta^{-6})d\theta = 50/3$$

$$EPV = E(200\theta^3) = \int_1^\infty 200\theta^3(5\theta^{-6})d\theta = 500$$

$$VHM = Var(10\theta^2) = \int_1^\infty (10\theta^2)^2(5\theta^{-6})d\theta - (50/3)^2 = 222.22$$

$$k = 500/222.22 = 2.25.$$

Question #12
Key: A

$c = \exp(.71(1) + .20(1)) = 2.4843$. Then $\hat{S}(t_0;\mathbf{z}) = \hat{S}_0(t_0)^c = .65^{2.4843} = .343$.

Question #13
Key: E

Y and X are linear combinations of the same two normal random variables, so they are bivariate normal. Thus $E(Y|X) = E(Y) + [\mathrm{Cov}(Y,X)/\mathrm{Var}(X)][X - E(X)]$. From the definitions of Y and X, $\mathrm{E}(Y) = a$, $E(X) = d$, $\mathrm{Var}(X) = e^2 + f^2$, and $\mathrm{Cov}(Y,X) = be + cf$.

Question #14
Key: D

$$\Pr(\theta=1 \mid X=5)=\frac{f(5\mid\theta=1)\Pr(\theta=1)}{f(5\mid\theta=1)\Pr(\theta=1)+f(5\mid\theta=3)\Pr(\theta=3)}$$
$$=\frac{(1/36)(1/2)}{(1/36)(1/2)+(3/64)(1/2)}=16/43$$
$$\Pr(X_2>8\mid X_1=5)=\Pr(X_2>8\mid\theta=1)\Pr(\theta=1\mid X_1=5)+\Pr(X_2>8\mid\theta=3)\Pr(\theta=3\mid X_1=5)$$
$$=(1/9)(16/43)+(3/11)(27/43)=.2126.$$

For the last line, $\Pr(X>8\mid\theta)=\int_8^\infty \theta(x+\theta)^{-2}\,dx=\theta(8+\theta)^{-1}$ is used.

Question #15
Key: C

The sample mean for X is 720 and for Y is 670. The mean of all 8 observations is 695.

$$\hat{v}=\frac{\begin{array}{c}(730-720)^2+(800-720)^2+(650-720)^2+(700-720)^2\\+(655-670)^2+(650-670)^2+(625-670)^2+(750-670)^2\end{array}}{2(4-1)}=3475$$
$$\hat{a}=\frac{(720-695)^2+(670-695)^2}{2-1}-\frac{3475}{4}=381.25$$
$$\hat{k}=3475/381.25=9.1148$$
$$\hat{Z}=\frac{4}{4+9.1148}=.305$$
$$P_c=.305(670)+.695(695)=687.4.$$

Question #16
Key: B

There are 430 observations. The expected counts are 430(.2744) = 117.99, 430(.3512) =151.02, 430(.3744) = 160.99. The test statistic is

$$\frac{(112-117.99)^2}{117.99}+\frac{(180-151.02)^2}{151.02}+\frac{(138-160.99)^2}{160.99}=9.15.$$

Question #17
Key: C

See pages 493-98 of the text.

Question #18
Key: B

From the information, the asymptotic variance of $\hat{\theta}$ is 1/4n. Then
$Var(2\hat{\theta}) = 4Var(\hat{\theta}) = 4(1/4n) = 1/n.$
Note that the delta method is not needed for this problem, although using it leads to the same answer.

Question #19
Key: A

The posterior probability density is

$$\pi(p\,|\,1,1,1,1,1,1,1,1) \propto \Pr(1,1,1,1,1,1,1,1\,|\,p)\pi(p) \propto p^8(2) \propto p^8.$$

$$\pi(p\,|\,1,1,1,1,1,1,1,1) = \frac{p^8}{\int_0^{.5} p^8 dp} = \frac{p^8}{(.5^9)/9} = 9(.5^{-9})p^8.$$

$$\Pr(X_9 = 1\,|\,1,1,1,1,1,1,1,1) = \int_0^{.5} \Pr(X_9 = 1\,|\,p)\pi(p\,|\,1,1,1,1,1,1,1,1)dp$$

$$= \int_0^{.5} p9(.5^{-9})p^8 dp = 9(.5^{-9})(.5^{10})/10 = .45.$$

Question #20
Key: A

The restricted model is $Y_t = X_t + \varepsilon_t$ and so $\hat{Y}_t = X_t$ and the ESS is $\sum_{t=1}^{5}(Y_t - X_t)^2 = 99{,}374.$ The F statistic is [(99,374 − 69,843)/2]/[69,843/3] = 0.6, which is less than the 95th percentile of the F distribution with 2 and 3 degrees of freedom.

Question #21
Key: A

${}_3\hat{p}_1 = \frac{18}{27}\frac{26}{32}\frac{20}{25} = \frac{13}{30}.$ Greenwood's approximation is

$$\left(\frac{13}{30}\right)^2\left(\frac{9}{18(27)} + \frac{6}{26(32)} + \frac{5}{20(25)}\right) = .0067.$$

Question #22
Key: D

$\hat{H}(3)$ = 5/30 + 9/27 + 6/32 = 0.6875
$V\hat{a}r(\hat{H}(3))$ = 5/(30)^2 + 9/(27)^2 + 6/(32)^2 = 0.02376

The 95% log-transformed confidence interval is:
$\hat{H}(3)U$, where $U = \exp\left(\pm\dfrac{1.96\sqrt{.02376}}{.6875}\right) = \exp(\pm 0.43945)$
The confidence interval is:
[0.6875 exp(–0.43945), 0.6875 exp(0.43945)] = [0.443, 1.067].

Question #23
Key: D

The means are .5(250) + .3(2,500) + .2(60,000) = 12,875 and .7(250) + .2(2,500) + .1(60,000) = 6,675 for risks 1 and 2 respectively.
The variances are .5(250)^2 + .3(2,500)^2 + .2(60,000)^2 – 12,875^2 = 556,140,625 and .7(250)^2 + .2(2,500)^2 + .1(60,000)^2 – 6,675^2 = 316,738,125 respectively.

The overall mean is (2/3)(12,875) + (1/3)(6,675) = 10,808.33 and so
EPV = (2/3)(556,140,625) + (1/3)(316,738,125) = 476,339,792 and
VHM = (2/3)(12,875)^2 + (1/3)(6,675)^2 – 10,808.33^2 = 8,542,222. Then,
k = 476,339,792/8,542,222 = 55.763 and Z = 1/(1 + 55.763) = .017617.
The credibility estimate is .017617(250) + .982383(10,808.33) = 10,622.

Question #24
Key: D

The first two sample moments are 15 and 500, and the first two population moments are $E(X) = .5(\theta+\sigma)$ and $E(X^2) = .5(2\theta^2 + 2\sigma^2) = \theta^2 + \sigma^2$. These can be obtained either through integration or by recognizing the density function as a two-point mixture of exponential densities. The equations to solve are $30 = \theta + \sigma$ and $500 = \theta^2 + \sigma^2$. From the first equation, $\sigma = 30 - \theta$ and substituting into the second equation gives $500 = \theta^2 + (30-\theta)^2 = 2\theta^2 - 60\theta + 900$. The quadratic equation has two solutions, 10 and 20. Because $\theta > \sigma$ the answer is 20.

Question #25
Key: E

To see that (E) is false, take expectations on both sides of the model:

$$E(y_t) = .8E(y_{t-1}) + 2 + E(\varepsilon_t) - .5E(\varepsilon_{t-1})$$

$$\mu = .8\mu + 2 + 0 - 0$$

$$\mu = 10$$

The other answers can be shown to be true be looking at various properties of an ARMA(1,1) model.

Question #26
Key: D

There are four possible samples, (5,5), (5,9), (9,5), and (9,9). For each, the estimator *g* must be calculated. The values are 0, 4, 4, and 0 respectively. Assuming a population in which the values 5 and 9 each occur with probability .5, the population variance is $.5(5-7)^2 + .5(9-7)^2 = 4$. The mean square error is approximated as $.25[(0-4)^2 + (4-4)^2 + (4-4)^2 + (0-4)^2] = 8$.

Question #27
Key: B

From the Poisson distribution, $\mu(\lambda) = \lambda$ and $v(\lambda) = \lambda$. Then, $\mu = E(\lambda) = 6/100 = .06$, $EPV = E(\lambda) = .06$, $VHM = Var(\lambda) = 6/100^2 = .0006$ where the various moments are evaluated from the gamma distribution. Then, $k = .06/.0006 = 100$ and $Z = 450/(450+100) = 9/11$ where the 450 is the total number of insureds contributing experience. The credibility estimate of the expected number of claims for one insured in month 4 is (9/11)(25/450) + (2/11)(.06) = .056364. For 300 insureds the expected number of claims is 300(.056364) = 16.9.

Question #28
Key: C

The likelihood function is $L(\alpha,\theta) = \prod_{j=1}^{200} \frac{\alpha\theta^\alpha}{(x_j+\theta)^{\alpha+1}}$ and its logarithm is $l(\alpha,\theta) = 200\ln(\alpha) + 200\alpha\ln(\theta) - (\alpha+1)\sum_{i=1}^{200}\ln(x_i+\theta)$. When evaluated at the hypothesized values of 1.5 and 7.8, the loglikelhood is –821.77. The test statistic is 2(821.77–817.92) = 7.7. With two degrees of freedom (0 free parameters in the null hypothesis versus 2 in the alternative), the test statistic falls between the 97.5th percentile (7.38) and the 99th percentile (9.21).

Question #29
Key: B

The least squares estimate is $\hat{\beta} = \sum X_i Y_i / \sum X_i^2$ and so the first residual and its variance are

$$\begin{aligned}\hat{\varepsilon}_1 &= Y_1 - \hat{\beta} X_1 = \beta X_1 + \varepsilon_1 - \frac{X_1 \sum X_i(\beta X_i + \varepsilon_i)}{\sum X_i^2} \\ &= \beta X_1 + \varepsilon_1 - \beta X_1 - X_1 \frac{\sum X_i \varepsilon_i}{\sum X_i^2} = \varepsilon_1 - \frac{\varepsilon_1 + 2\varepsilon_2 + 3\varepsilon_3}{14} \\ &= \frac{13\varepsilon_1 - 2\varepsilon_2 - 3\varepsilon_3}{14}.\end{aligned}$$

$$Var(\hat{\varepsilon}_1) = \frac{169(1) + 4(9) + 9(16)}{196} = \frac{349}{196} = 1.78.$$

Question #30
Key: E

Assume that $\theta > 5$. Then the expected counts for the three intervals are $15(2/\theta) = 30/\theta, 15(3/\theta) = 45/\theta,$ and $15(\theta - 5)/\theta = 15 - 75/\theta$ respectively. The quantity to minimize is

$$\frac{1}{5}\left[(30\theta^{-1} - 5)^2 + (45\theta^{-1} - 5)^2 + (15 - 75\theta^{-1} - 5)^2\right].$$

Differentiating (and ignoring the coefficient of 1/5) gives the equation $-2(30\theta^{-1} - 5)30\theta^{-2} - 2(45\theta^{-1} - 5)45\theta^{-2} + 2(10 - 75\theta^{-1})75\theta^{-2} = 0$. Multiplying through by θ^3 and dividing by 2 reduces the equation to $-(30 - 5\theta)30 - (45 - 5\theta)45 + (10\theta - 75)75 = -8550 + 1125\theta = 0$ for a solution of $\hat{\theta} = 8550/1125 = 7.6$.

Question #31
Key: E

$\pi(\theta \mid 1) \propto \theta(1.5\theta^{.5}) \propto \theta^{1.5}$. The required constant is the reciprocal of $\int_0^1 \theta^{1.5} d\theta = \theta^{2.5}/2.5\Big|_0^1 = .4$ and so $\pi(\theta \mid 1) = 2.5\theta^{1.5}$. The requested probability is

$$\Pr(\theta > .6 \mid 1) = \int_{.6}^1 2.5\theta^{1.5} d\theta = \theta^{2.5}\Big|_{.6}^1 = 1 - .6^{2.5} = .721.$$

Question #32
Key: A

k	kn_k / n_{k-1}
0	
1	0.81
2	0.92
3	1.75
4	2.29
5	2.50
6	3.00

Positive slope implies that the negative binomial distribution is a good choice. Alternatively, the sample mean and variance are 1.2262 and 1.9131 respectively. With the variance substantially exceeding the mean, the negative binomial model is again supported.

Question #33
Key: B

The required equation is

$$(1-.7B)(1-B)(1+\psi_1 B+\cdots)=1+.3B$$

$$1+(\psi_1-1.7)B+\cdots=1+.3B$$

$$\psi-1.7=.3, \quad \psi_1=2.$$

The variance of the forecast error two steps ahead is $\sigma_\varepsilon^2(1+\psi_1^2)=1(1+2^2)=5.$

Question #34
Key: B

The likelihood function is $\frac{e^{-1/(2\theta)}}{2\theta}\cdot\frac{e^{-2/(2\theta)}}{2\theta}\cdot\frac{e^{-3/(2\theta)}}{2\theta}\cdot\frac{e^{-15/(3\theta)}}{3\theta}=\frac{e^{-8/\theta}}{24\theta^4}$. The loglikelihood function is $-\ln 24-4\ln(\theta)-8/\theta$. Differentiating with respect to θ and setting the result equal to 0 yields $-\frac{4}{\theta}+\frac{8}{\theta^2}=0$ which produces $\hat{\theta}=2$.

Question #35
Key: E

The absolute difference of the credibility estimate from its expected value is to be less than or equal to $k\mu$ (with probability *P*). That is,

$$\left|[ZX_{partial} + (1-Z)M] - [Z\mu + (1-Z)M]\right| \le k\mu$$

$$-k\mu \le ZX_{partial} - Z\mu \le k\mu.$$

Adding μ to all three sides produces answer choice (E).

Question #36
Key: C

$\hat{\sigma}^2 = 282.82/(15-4) = 25.71$. $Var(\hat{\beta}_3 - \hat{\beta}_2) = Var(\hat{\beta}_3) + Var(\hat{\beta}_2) - 2Cov(\hat{\beta}_3, \hat{\beta}_2)$ which can be estimated as 25.71[2.14 + 0.03 – 2(0.11)] = 50.14. The standard error is the square root of the estimated variance, 7.1.

Question #37
Key: C

In general,

$$E(X^2) - E[(X \wedge 150)^2] = \int_0^{200} x^2 f(x)dx - \int_0^{150} x^2 f(x)dx - 150^2 \int_{150}^{200} f(x)dx = \int_{150}^{200} (x^2 - 150^2) f(x)dx.$$

Assuming a uniform distribution, the density function over the interval from 100 to 200 is 6/7400 (the probability of 6/74 assigned to the interval divided by the width of the interval). The answer is

$$\int_{150}^{200} (x^2 - 150^2)\frac{6}{7400}dx = \left(\frac{x^3}{3} - 150^2 x\right)\frac{6}{7400}\Bigg|_{150}^{200} = 337.84.$$

Question #38
Key: C

See pages 476-477.

Question #39
Key: B

The probabilities are from a binomial distribution with 6 trials. Three successes were observed.

$$\Pr(3\,|\,\mathrm{I}) = \binom{6}{3}(.1)^3(.9)^3 = .01458, \Pr(3\,|\,\mathrm{II}) = \binom{6}{3}(.2)^3(.8)^3 = .08192,$$

$$\Pr(3\,|\,\mathrm{III}) = \binom{6}{3}(.4)^3(.6)^3 = .27648$$

The probability of observing three successes is .7(.01458) + .2(.08192) + .1(.27648) = .054238. The three posterior probabilities are:

$$\Pr(\mathrm{I}\,|\,3) = \frac{.7(.01458)}{.054238} = .18817,\ \Pr(\mathrm{II}\,|\,3) = \frac{.2(.08192)}{.054238} = .30208,\ \Pr(\mathrm{III}\,|\,3) = \frac{.1(.27648)}{.054238} = .50975.$$

The posterior probability of a claim is then
.1(.18817) + .2(.30208) + .4(.50975) = .28313.

Question #40
Key: E

$.542 = \hat{F}(n) = 1 - e^{-\hat{H}(n)}, \quad \hat{H}(n) = .78.$ The Nelson-Aalen estimate is the sum of successive *s*/*r* values. From the problem statement, *r* = 100 at all surrender times while the *s*-values follow the pattern 1, 2, 3, …. Then,

$.78 = \frac{1}{100} + \frac{2}{100} + \cdots + \frac{n}{100} = \frac{n(n+1)}{200}$ and the solution is *n* = 12.