A.1 Lifetime distributions

1. Give lots of examples of settings in which we might model "lifetimes" using the framework of survival analysis.

2. (a) Let $L_1, \ldots, L_n$ be independent $\mathrm{Exp}(\lambda)$ random variables. Show that the maximum likelihood estimator for λ is given by
$$\hat{\lambda} = \frac{n}{L_1 + \cdots + L_n}.$$

   (b) The following data resulted from a life test of refrigerator motors (hours to burnout):

Hours to burnout				
104.3	158.7	193.7	201.3	206.2
227.8	249.1	307.8	311.5	329.6
358.5	364.3	370.4	380.5	394.6
426.2	434.1	552.6	594.0	691.5

   i. Assuming refrigerator motors have $\mathrm{Exp}(\lambda)$ lifetimes, determine the maximum likelihood estimate for λ.

   ii. Still assuming $\mathrm{Exp}(\lambda)$ lifetimes, calculate the Fisher information and construct approximate 95% confidence intervals for λ and $1/\lambda$ using the approximate Normal distribution of the maximum likelihood estimator.

   iii. Still assuming $\mathrm{Exp}(\lambda)$ lifetimes, show that $2n\lambda/\hat{\lambda} \sim \chi^2_{2n}$. Let a be such that $\mathbb{P}(2n\lambda/\hat{\lambda} \le a) = \alpha/2$ and b such that $\mathbb{P}(2n\lambda/\hat{\lambda} \ge b) = \alpha/2$. Deduce an exact 95% confidence interval for $1/\lambda$.

   iv. Produce a histogram of the data and comment.

   v. Merge columns of your histogram appropriately to test whether the hypothesis of $\mathrm{Exp}(\lambda)$ lifetimes can be rejected. Use a χ^2 goodness of fit test.

3. (a) Let $T_1, \ldots, T_m$ be independent continuous nonnegative random variables with hazard functions $h_1(\cdot), \ldots, h_m(\cdot)$. Prove that $T = \min(T_1, \ldots, T_m)$ has hazard function $h_1(\cdot) + \cdots + h_m(\cdot)$.

   (b) A Weibull distribution with rate k and exponent n has hazard rate kt^n. Let $T_1, \ldots T_m$ be independent Weibull random variables with rate parameters $k_1, \ldots k_m$ and with common exponent n. Find the distribution of $T = \min(T_1, \ldots, T_m)$.

   (c) A truncated exponential distribution with parameter λ and maximal age ω has density proportional to $\lambda e^{-\lambda t}$ on $[0, \omega]$ and 0 elsewhere. Calcluate the hazard function of the distribution. Find the limit in distribution as $\lambda \downarrow 0$.

4. Suppose that lifetimes are exponentially distributed with rate μ, and that we have a prior distribution on μ which is a gamma distribution with shape parameter α and rate parameter β; that is,
$$f_\mu(m) = \frac{\beta^\alpha m^{\alpha-1} e^{-\beta m}}{\Gamma(\alpha)}.$$
Suppose we observe lifetimes $T_1, T_2, \ldots, T_n$. Show that the posterior distribution on μ is also a gamma distribution, and give its parameters.

5. We can obtain a class of distributions known as *exponential mixtures* by replacing the rate parameter of the exponential distribution by a positive random variable M, which may be discrete or continuous.

   The distribution of T conditional on M is given by

   $$f_{T|M=\lambda}(t) = \lambda e^{-\lambda t},$$

   so that the density of T is given by

   $$f_T(t) = \int_0^\infty \lambda e^{-\lambda t} f_M(\lambda) d\lambda \quad \text{or} \quad f_T(t) = \sum_\lambda \lambda e^{-\lambda t} \mathbb{P}(M = \lambda)$$

   in the continuous case or discrete case respectively.

   (a) Show that T has mean and variance given by

   $$\mathbb{E}(T) = \mathbb{E}\left(\frac{1}{M}\right) \qquad \text{and} \qquad \text{Var}(T) = 2\mathbb{E}\left(\frac{1}{M^2}\right) - \left(\mathbb{E}\left(\frac{1}{M}\right)\right)^2,$$

   and survival function

   $$\overline{F}_T(t) = \mathcal{M}_M(-t), \qquad \text{where } \mathcal{M}_M(c) = \mathbb{E}(e^{cM})$$

   is the moment generating function of M.

   (b) Consider the distribution with hazard function

   $$h(t) = \rho_0 + \rho_1 e^{\rho_2 t}$$

   where $\rho_0, \rho_1 > 0$. (If $\rho_2 > 0$, this is known as a Gompertz-Makeham distribution). Show that this distribution can be obtained as an exponential mixture provided $\rho_2 \leq 0$, and determine the distribution of the mixing random variable M.
   Hint: Calculate the moment generating function of a Poi(ν) *random variable* $\tilde{M}$ *and adjust as necessary.*

6. Suppose we observe ℓ_0 independent and identically distributed lifetimes and consider the random variables behind associated lifetable entries d_x and ℓ_x, $x \geq 0$.

   (a) Show that $\mathbb{E}(d_x - q_x\ell_x) = 0$ and $\text{Var}(d_x - q_x\ell_x) = q_x(1 - q_x)\mathbb{E}(\ell_x)$.
   Hint: Condition on ℓ_x. *What is the conditional distribution of* d_x *given* ℓ_x?

   (b) Show that the MLE for q_0 is $\hat{q}_0 = d_0/\ell_0$. Is it unbiased? What is its variance? What is the MLE $\hat{q}_x$ for $x \geq 1$? Is it unbiased? (Pay particular attention to different values that ℓ_x may take.) Calculate the relevant Fisher information matrix, and use it to give approximations to the variances of $\hat{q}_x$ for large ℓ_0, and estimates for these variances induced by the observed values of $\hat{q}_x$.