

There are 2 questions, you need to answer both. Do not blindly copy from your cheat sheet, your cheat sheet must be submitted with your solutions.

Good Luck and have a great summer

- (1) Let us suppose that T and C are exponentially distributed random variables, where the density of T is $\frac{1}{\lambda} \exp(-t/\lambda)$ and the density of C is $\frac{1}{\mu} \exp(-c/\mu)$. [25]

- (i) Evaluate the probability $P(T < C + x)$, where x is some finite constant.
- (ii) Let us suppose that $\{T_i\}_i$ and $\{C_i\}_i$ are iid survival and censoring times (T_i and C_i are independent of each other), where the densities of T_i and C_i are $f_T(t; \lambda) = \frac{1}{\lambda} \exp(-t/\lambda)$ and $f_C(c; \mu) = \frac{1}{\mu} \exp(-c/\mu)$ respectively. Let $Y_i = \min(T_i, C_i)$ and $\delta_i = 1$ if $Y_i = T_i$ and zero otherwise. Suppose λ and μ are unknown. We use the following likelihood to estimate λ

$$\mathcal{L}_n(\lambda) = \sum_{i=1}^T \delta_i \log f_T(Y_i; \lambda) + \sum_{i=1}^T (1 - \delta_i) \log \mathcal{F}_T(Y_i; \lambda),$$

where \mathcal{F}_T is the survival function.

Let $\hat{\lambda}_n = \arg \max \mathcal{L}_n(\lambda)$. Show that $\hat{\lambda}_n$ is an asymptotically, biased estimator of λ (you can assume that $\hat{\lambda}_n$ converges to some constant).

- (iii) Based on your results in parts (i) and (ii) construct estimators of λ and μ which are asymptotically unbiased.
- (2) Let us suppose that $\mathcal{F}_1(t)$ and $\mathcal{F}_2(t)$ are two survival functions. Let x denote a univariate regressor. [25]
- (i) Show that $\mathcal{F}(t; x) = p\mathcal{F}_1(t)^{\exp(\beta_1 x)} + (1-p)\mathcal{F}_2(t)^{\exp(\beta_2 x)}$ is a valid survival function and obtain the corresponding density function.
- (ii) Suppose that T_i are survival times and x_i is a univariate regressor which exerts an influence on T_i . Let $Y_i = \min(T_i, c)$, where c is a common censoring time. $\{T_i\}$ are independent random variables with survival function $\mathcal{F}(t; x_i) = p\mathcal{F}_1(t)^{\exp(\beta_1 x_i)} + (1-p)\mathcal{F}_2(t)^{\exp(\beta_2 x_i)}$, where both \mathcal{F}_1 and \mathcal{F}_2 are known, but p , β_1 and β_2 are unknown.

State the censored likelihood and show that the EM-algorithm together with iterative least squares in the maximisation step can be used to maximise this likelihood (sufficient details need to be given such that your algorithm can be easily coded).