Ph.D. Prelim: Exam A
Distribution Theory & Regression Analysis

January 15, 2016

1. Let X_1 and X_2 be independent non-negative continuous random variables with densities f_1 and f_2 and distribution functions F_1 and F_2 respectively. Show that

$$P(X_1 < X_2 | \min(X_1, X_2) = t) = \frac{\lambda_1(t)}{\lambda_1(t) + \lambda_2(t)},$$

where $\lambda_i(t)$ is the failure rate of X_i, defined as

$$\lambda_i(t) = f_i(t)/(1 - F_i(t)), \quad i = 1, 2.$$

2. Let X_1, X_2, X_3 be independent and distributed uniformly on $(0, 1)$. Let $X_{(1)}, X_{(2)}, X_{(3)}$ be the corresponding order statistics. Find the density of the pair

$$\left(\frac{X_{(1)}}{X_{(2)}}, \frac{X_{(2)}}{X_{(3)}} \right).$$

Are the ratios independent? Justify your answer.

3. This problem has two independent parts.
   (a) Let $Y_1, \ldots, Y_n$ be independent exponential random variables with respective rates $\lambda_1, \ldots, \lambda_n$. Compute $P(Y_1 \le \min(Y_2, \ldots, Y_n))$.
   (b) Suppose that for each $t \ge 0$, $N(t)$ has a Poisson distribution with parameter λt, where $\lambda > 0$ is a constant. Let T be a nonnegative random variable with mean μ and variance σ^2 independent of $\{N(t), t \ge 0\}$. Compute $\mathrm{Cov}(T, N(T))$ and $\mathrm{Var}(N(T))$.

4. Consider the linear model $Y_i = \beta X_i + \epsilon_i$, where $E(\epsilon_i) = 0$ and $\mathrm{Var}(\epsilon_i) = \sigma^2$ for $i = 1, 2, \ldots, n$ with β and X_i scalars. We assume that Y_i's are independent. The X_i's are not random.
   (a) Write out explicitly in terms of the Ys and Xs; what the least squares estimator of β is (call it $\hat{\beta}$) and give an explicit expression for $\mathrm{Var}(\hat{\beta})$. Also give an expression for $\hat{\sigma}^2$ and show that it is an unbiased estimator of σ^2.
   (b) An alternate estimator arises by considering $Z_i = Y_i/X_i$.
      i. Show that $\bar{Z} = \sum_{i=1}^{n} Z_i/n$ is an unbiased estimator of β.
      ii. Find the variance of $\bar{Z}$ and show directly that this variance is greater than or equal to the variance of $\hat{\beta}$.

5. Let $E_1, \ldots, E_k$ be k events, then we have the Bonferroni inequality:

$$\Pr(\bigcap_{i=1}^{k} E_i) \geq 1 - \sum_{i=1}^{k} \Pr(E_i^C).$$

We want to apply this inequality to construct simultaneous confidence intervals for k contrasts of means. Consider independent Y_{ij} for $j = 1, 2, \ldots, n_i$ and $i = 1, 2, 3$, where $Y_{ij} \sim N(\mu_i, \sigma^2)$.

(a) Prove the Bonferroni inequality.

(b) Suppose we want to construct simultaneous confidence intervals for μ_1 and μ_2 with confidence coefficient 0.90. How shall we construct such confidence intervals? Please justify your answers using the Bonferroni inequality.

(c) Describe how to construct simultaneous confidence intervals for $\mu_2 - \mu_1$ and $\frac{1}{2}(\mu_2 + \mu_3) - \mu_1$ with confidence at least 0.95. (Note: you need to write down explicit formulae for your notations.)

6. In a normal error simple linear regression model, random errors are assumed to be i.i.d. (identically independently distributed).

(a) What is the unbiased estimator of σ^2 we usually use? Prove its unbiasedness. What is the MLE of the σ^2? Show that the MLE is biased.

(b) What is R^2 and how is it calculated? How is R^2 related with the Pearson correlation coefficient defined as

$$r = \frac{\sum_{i=1}^{n}(X_i - \bar{X})(Y_i - \bar{Y})}{\sqrt{\sum_{i=1}^{n}(X_i - \bar{X})^2 \sum_{i=1}^{n}(Y_i - \bar{Y})^2}}?$$

Please justify your answer mathematically.