STA 302f13 Assignment Two[1]

These problems are preparation for the quiz in tutorial on Friday September 27th, and are not to be handed in. Starting with Problem 2, you can play a little game. Try not to do the same work twice. Instead, use results of earlier problems whenever possible.

1. Sometimes, you want the least squares line to go through the origin, so that predicted Y automatically equals zero when $x = 0$. For example, suppose the cases are half-kilogram batches of rice purchased from grocery stores. The independent variable x is concentration of arsenic in the rice before washing, and the dependent variable Y is concentration of arsenic after washing. Discounting the very unlikely possibility that arsenic contamination can happen *during* washing, you want to use your knowledge that zero arsenic before washing implies zero arsenic after washing. You will use your knowledge by building it into the statistical model.

   Accordingly, let $Y_i = \beta x_i + \epsilon_i$ for $i = 1, \ldots, n$, where $\epsilon_1, \ldots, \epsilon_n$ are a random sample from a distribution with expected value zero and variance σ^2, and β and σ^2 are unknown constants. The numbers $x_1, \ldots, x_n$ are known, observed constants.

   (a) What is $E(Y_i)$?

   (b) What is $Var(Y_i)$?

   (c) Find the Least Squares estimate of β by minimizing the function

   $$Q(\beta) = \sum_{i=1}^{n} (Y_i - \beta x_i)^2$$

   over all values of β. Let $\widehat{\beta}$ denote the point at which $Q(\beta)$ is minimal.

   (d) Give the equation of the least-squares line. Of course it's the *constrained* least-squares line, passing through $(0, 0)$.

   (e) Recall that a statistic is an *unbiased estimator* of a parameter if the expected value of the statistic is equal to the parameter. Is $\widehat{\beta}$ an unbiased estimator of β? Answer Yes or No and show your work.

   (f) Let $\widehat{\beta}_2 = \frac{\overline{Y}_n}{\overline{x}_n}$. Is $\widehat{\beta}_2$ also unbiased for β? Answer Yes or No and show your work.

   (g) *This last part is a challenge for your entertainment. It will not be on the quiz or the final exam.* Prove that $\widehat{\beta}$ is a more accurate estimator than $\widehat{\beta}_2$ in the sense that it has smaller variance. Hint: The sample variance of the independent variable values cannot be negative.

2. Denote the moment-generating function of a random variable Y by $M_Y(t)$. The moment-generating function is defined by $M_Y(t) = E(e^{Yt})$.

   (a) Let a be a constant. Prove that $M_{aX}(t) = M_X(at)$.

   (b) Prove that $M_{X+a}(t) = e^{at} M_X(t)$.

   (c) Let $X_1, \ldots, X_n$ be *independent* random variables. Prove that

   $$M_{\sum_{i=1}^{n} X_i}(t) = \prod_{i=1}^{n} M_{X_i}(t).$$

   For convenience, you may assume that $X_1, \ldots, X_n$ are all continuous, so you will integrate.

[1] Copyright information is at the end of the last page.

3. Recall that if $X \sim N(\mu, \sigma^2)$, it has moment-generating function $M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2}$.

   (a) Let $X \sim N(\mu, \sigma^2)$ and $Y = aX + b$, where a and b are constants. Find the distribution of Y. Show your work.

   (b) Let $X \sim N(\mu, \sigma^2)$ and $Z = \frac{X-\mu}{\sigma}$. Find the distribution of Z.

   (c) Let $X_1, \ldots, X_n$ be random sample from a $N(\mu, \sigma^2)$ distribution. Find the distribution of $Y = \sum_{i=1}^n X_i$.

   (d) Let $X_1, \ldots, X_n$ be random sample from a $N(\mu, \sigma^2)$ distribution. Find the distribution of the sample mean $\overline{X}$.

   (e) Let $X_1, \ldots, X_n$ be random sample from a $N(\mu, \sigma^2)$ distribution. Find the distribution of $Z = \frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}$.

4. A Chi-squared random variable X with parameter $\nu > 0$ has moment-generating function $M_X(t) = (1-2t)^{-\nu/2}$.

   (a) Let $X_1, \ldots, X_n$ be independent random variables with $X_i \sim \chi^2(\nu_i)$ for $i = 1, \ldots, n$. Find the distribution of $Y = \sum_{i=1}^n X_i$.

   (b) Let $Z \sim N(0,1)$. Find the distribution of $Y = Z^2$.

   (c) Let $X_1, \ldots, X_n$ be random sample from a $N(\mu, \sigma^2)$ distribution. Find the distribution of $Y = \frac{1}{\sigma^2}\sum_{i=1}^n (X_i - \mu)^2$.

   (d) Let $Y = X_1 + X_2$, where X_1 and X_2 are independent, $X_1 \sim \chi^2(\nu_1)$ and $Y \sim \chi^2(\nu_1 + \nu_2)$, where ν_1 and ν_2 are both positive. Show $X_2 \sim \chi^2(\nu_2)$.

   (e) Let $X_1, \ldots, X_n$ be random sample from a $N(\mu, \sigma^2)$ distribution. Show

$$\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1),$$

   where $S^2 = \frac{\sum_{i=1}^n (X_i - \overline{X})^2}{n-1}$. Hint: $\sum_{i=1}^n (X_i - \mu)^2 = \sum_{i=1}^n \left(X_i - \overline{X} + \overline{X} - \mu\right)^2 = \ldots$
   For this question, you may use the independence of $\overline{X}$ and S^2 without proof.
