Statistics & Probability (191506103)

Final Exam (with 7 questions, 1 formula-sheet and 3 tables)

Monday 28/10/19, 08:45 − 11:45

Full Marks : 35

1. Derive the probability density function (pdf) of a $\chi^2_{(1)}$ random variable. In other words, suppose Z is a standard normal random variable, with pdf $\quad \phi(z) = \frac{1}{\sqrt{2\pi}} e^{-z^2/2}, \quad -\infty < z < +\infty$. Define the new r.v. $X = Z^2$, and obtain the pdf of X. [4]

2. Suppose the random variables X and Y are (jointly) uniformly distributed over the triangle determined by the points $(0,0)$, $(0,1)$ and $(\frac{1}{2}, \frac{1}{2})$. In other words, the joint pdf is given by

$$f(x, y) = 4, \quad 0 \le x \le y; \quad x + y \le 1 \quad \text{and} \quad = 0, \quad \text{elsewhere.}$$

   (a.) Find the (marginal) pdf of Y. [2]

   (b.) Find the conditional pdf, $f_{X|Y=y}(\cdot)$ of X given $Y = y$. [2]

   (c.) Recall that $E(X|Y)$ is a random variable, which is a function of Y. Express the random variable in terms of Y. [2]

3. For a certain kind of flower bulb it is known that 3% of the bulbs do not germinate. The bulbs are packaged and sold in boxes of 10 with the guarantee that at least 9 of the 10 bulbs will germinate.

   (a.) What is the probability that an arbitrary box will not have the guaranteed property? [2]

   [If you do not find the answer in (a.)*, assume it to be* 0.04 *for the following part.]*

   (b.) A big shipment of 1000 boxes was received by a dealer. What is the probability that at most 30 of these boxes will fail to satisfy the guarantee? Give as accurate an answer as you can from the materials provided. [4]

   [Answer obtained directly from an advanced/graphical calculator would not be accepted.]

4. Suppose X and Y are independent normal random variables with expectations 0 and 2 and variances 1 and 4, respectively. Define a new random variable $Z := 2X + 3Y$. Use the moment generating function (mgf) technique to find the probability distribution of Z. Recall that the mgf for $N(\mu, \sigma^2)$ is given by $m(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$. [Motivate your answer clearly.] [3]

5. Consider two independent observations X and Y from two Binomial populations with unknown (and possibly different) success probabilities p_1 and p_2. In other words, X and Y are independent with $X \sim Bin(n_1, p_1)$ and $Y \sim Bin(n_2, p_2)$. Suppose our interest lies in $p_1 - p_2$, the difference between two probabilities.

   (a.) Show that $\dfrac{X}{n_1} - \dfrac{Y}{n_2}$ is an unbiased estimator for $p_1 - p_2$. [2]

   (b.) Calculate the mean squared error (MSE) of the estimator in part (a.). [2]

   [Hint: Recall that MSE = Variance + Bias^2.]

6. The melting point of hydrogenated vegetable oil is assumed to be normally distributed. A new brand of hydrogenated vegetable oil has arrived in the market. We would like to estimate the (expected) melting point of the new brand. To this end, the melting points of 16 samples from the new brand were determined. This resulted in a sample mean of 94.32 and a sample standard deviation of 1.275.

   (a.) Construct a 95% confidence interval for the true expected melting point for this new brand of oil. [3]

   (b.) If you calculate a 97% confidence interval (instead of 95%), based on the same data as give above, would that be narrower or wider than the one obtained in part (a.)? What would be your answer if you had calculated the (97%) confidence interval based on another data-set? Explain both of your answers without doing any actual calculation. [1+1]

7. How popular is the Tinder app for dating among the youths? A researcher has conducted a survey among 200 university students to find that 113 of the surveyed students have used the app.

   (a.) Can the researcher conclude from the survey that more than half of the university students use the Tinder app for dating? Help the researcher by performing an appropriate statistical hypothesis test using a level of significance of $\alpha = 0.05$. [5]

   (b.) Calculate the P-value of the obtained data according to the test procedure you used in part (a.). Would your answer to the question in part (a.) change, if you had used a significance level of 1% instead of 5%? Explain your answer using the P-value. [1+1]

Final grade $= \left\{ \dfrac{\text{score on exam}}{35} \times 9 + 1 \right\}$ (rounded off to an integer)