Course Project of Probability and Statistics

2019.10.27

Problem 1

Recall that the normal distribution $\mathcal{N}(x;\mu,\sigma^2)$ has pdf

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} \mathrm{e}^{-\frac{(x-\mu)^2}{2\sigma^2}}.$$

Suppose X is a continue random variable that has normal distribution.

1) Prove that $f(x;\mu,\sigma^2)$ is a legitimate pdf.

2) Compute the mean $E(x)$ and variance $V(x)$ of X.

The density curve corresponding to any normal distribution is bell-shaped and therefore symmetric. There are many practical situations in which the variable of interest to an investigator might have a skewed distribution. One family of distributions that has this property is the gamma family. A continuous random variable X is said to have a gamma distribution $\Gamma(x;\ \alpha,\beta)$ if the pdf of X is

$$f(x\,;\alpha\,,\beta) = \begin{cases} \dfrac{1}{\beta^\alpha \Gamma(\alpha)}\, x^{\alpha-1} e^{\frac{-x}{\beta}} & x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

where the parameters α and β satisfy $\alpha > 0$, $\beta > 0$.

3) Prove $\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$

4) Prove $f(x\,;\alpha\,,\beta)$ is a legitimate pdf.

5) The normal distribution and the gamma distribution have some kind of connection. Suppose X is a continue rv that has a normal distribution $\mathcal{N}(x;\ 0,1)$, compute the pdf of the continue rv X^2.

Two random variables X and Y are said to be independent if for every pair of x and y values,

$$f(x,y) = f_X(x)\cdot f_Y(y).$$

6) Suppose X and Y are independent, and both X and Y have normal distribution $\mathcal{N}(x;\ 0,1)$, prove that the random variables X^2+Y^2 and $\frac{X}{Y}$ are also independent.

Problem 2

In an experiment, we need to measure the voltage of an electric signal. Suppose the true voltage is μ. But every time when we measure the voltage, due to the existence of measurement errors, the actual observed value x is

$$x = \mu + \epsilon,$$

where ϵ represents the measurement error, and is often assumed following the standard normal distribution $p(\epsilon) = \mathcal{N}(x;\ 0, 1)$. Intuitively, we can infer the possible values of μ based on the observation x.

Because we don't have any knowledges about the true voltage values μ in advance, we assume it follows a normal distribution with zero mean and very large variance, that is, $p(\mu) = \mathcal{N}(\mu; 0, \sigma^2)$ (σ^2 could be set to a very large value, like 100, 10000 etc.).

Based on the backgrounds introduced above, please answer the following questions.

1) What is the conditional probability distribution $p(x|\mu)$ and the joint probability distribution $p(x, \mu)$?

2) We take a measurement on the voltage and denote the observed value as x_1. Given the observation x_1, what can we say about the probability distribution of μ, that is, $p(\mu|x_1)$?

3) Now, we take another measurement and denote the second observed value as x_2. During different measurements, we assume the voltage value keeps constant. Given the two observations x_1 and x_2, what can we say about the probability distribution of μ?

4) If there are n measurement values $x_1, x_2, \cdots, x_n$ observed, given these observations, what can we say about the probability distribution of μ?

5) Suppose we take a sequence of 20 measurements and obtain the following values: 1.74, 3.37, 2.64, 3.86, 3.00, 1.29, 3.65, 3.50, 2.73, 2.88, 3.13, 2.29, 2.66, 0.61, 2.03, 3.62, 4.61, 2.50, 3.98, 3.69.

   If σ^2 is set to 100, plot the probability distribution of μ when the first one, firth three, first five, first ten, all twenty measurement values shown above are given;

   Observing how these probability distributions evolve as more measurement values are given, discuss implications of the obtained results.

6) When σ^2 is set to values 0.1, 1, 10, please plot the corresponding probability distributions of μ, respectively (all the twenty measurements are given). Then, discuss the impacts of σ^2.

Requirements:

1) Choose one problem above, and finish the project independently. If any two reports are found to be identical, both reports will get score ZERO

2) Submit your report to the FTP address (ftp://222.200.180.156//郑培嘉老师/作业上传/中期考核 2019 年概率论与数理统计/) under the filename: <your student No.+your name>.rar or .zip

3) Deadline: 11:59PM, Nov. 10, 2019