CSCI 2244 – Homework 6

Out: Friday, October 11, 2019
Due: Friday, October 25, 2019, 11:59pm

This homework consists of **only written** exercises. You *must* type your solutions. See the "Assignments" section in the syllabus for advice about doing this. You should submit your homework via Canvas. In particular, you should upload a PDF file called:

`FirstName_LastName_Homework6.pdf`

Please use your full first name and last name, as they appear in official university records. The reason for doing so is that the TAs and I must match up these names with the entries in the gradebook.

1 Transforming Continuous Random Variables

Let X be a continuous random variable, and let $g : \mathbb{R} \to \mathbb{R}$ be a function. Set Y to be the random variable $g(X)$. If X has a CDF F_X and density f_x, we might wonder what the CDF and density of Y are. Section 4.2.4 of the textbook describes how to find the density f_Y when X is a uniform $[0, 1]$ random variable and g is a differentiable one-to-one function. Take a look at that.

What about when X is a general continuous random variable, not just a uniform one? In that case, we can work this out if g is a strictly increasing function with an inverse, g^{-1}. Then we have that:

$$\begin{aligned} F_Y(t) &= P(Y \leq t) && (1) \\ &= P(g(X) \leq t) && (2) \\ &= P(X \leq g^{-1}(t)) && (3) \\ &= F_X(g^{-1}(t)) && (4) \end{aligned}$$

Differentiating F_Y, we get then that:

$$f_Y(t) = f_X(g^{-1}(t)) \frac{d}{dt} \left(g^{-1}(t)\right) \tag{5}$$

Task 1.1 (12 pts). Let X be random variable with uniform distribution on the interval $[0, 1]$. Using the above facts or the result from the textbook, find the density and cumulative distribution of the following random variables:

(a) $Y = X + 2$

(b) $Z = X^3$

(c) $R = \log(X + 1)$

2 Expected Values and Variances of Continuous Random Variables

If X is a continuous random variable with density f_X, then the expected value of X can be found by:

$$\mathbb{E}[X] = \int_{-\infty}^{\infty} t f_X(t) dt$$

If $g : \mathbb{R} \to \mathbb{R}$ is a continuous function, then we have:

$$\mathbb{E}[g(X)] = \int_{\infty}^{\infty} g(t) f_X(t) dt$$

Task 2.1 (12 pts). Let X be a random variable with density f_X. Assume that $f_X(t) = 0$ if $t < -1$ or $t > 1$. Calculate $\mathbb{E}[X]$ and $\mathsf{Var}[X]$ for each case below, assuming that f_X has the value stated for $-1 < t < 1$:

(a) $f_X(t) = 1/2$

(b) $f_X(t) = |t|$

(c) $f_X(t) = \frac{3}{4}\left(1 - t^2\right)$

3 Convolution

Let X and Y be two independent continuous random variables with density functions f_X and f_Y respectively. Let $Z = X + Y$. Then Z has density f_Z given by the convolution of f_X and f_Y:

$$\begin{aligned} f_Z(t) = (f_X * f_Y)(t) &= \int_{-\infty}^{\infty} f_X(t-y) f_Y(y) dy \\ &= \int_{-\infty}^{\infty} f_Y(t-x) f_X(x) dx \end{aligned}$$

Task 3.1 (12 pts). Let X and Y be independent, with densities f_X and f_Y. Let $Z = X + Y$. Find the density f_Z of Z when f_X and f_Y have the following values:

(a) $f_X(t) = f_Y(t) = \begin{cases} \frac{t}{2} & \text{if } 0 < t < 2 \\ 0 & \text{otherwise} \end{cases}$

(b) $f_X(t) = \begin{cases} \lambda e^{-\lambda t} & \text{if } t > 0 \\ 0 & \text{otherwise} \end{cases}$

$f_Y(t) = \begin{cases} \mu e^{-\mu t} & \text{if } t > 0 \\ 0 & \text{otherwise} \end{cases}$

where $\mu > 0$, $\lambda > 0$ and $\mu \neq \lambda$

4 Normal Approximation

Task 4.1 (6 pts). Let $X_1, \ldots, X_n$ be independent $\mathsf{Bernoulli}(.5)$ random variables. Set $X = X_1 + \cdots + X_n$. By the Central Limit Theorem, we know that if n is large enough, then the distribution of X is approximately equal to a normal distribution with mean $\mathbb{E}[X]$ and standard deviation $\sqrt{\mathsf{Var}[X]}$. Consider the case when $n = 100$. Approximate $P(|X - 50| \geq z)$ for $z \in \{10, 20, 30, 40\}$ using this normal approximation. How do these values compare with the bounds we got from Chebyshev/Chernoff in homework 4?

In Python 3, you can compute the CDF of a normal distribution using `norm.cdf` from `scipy.stats` as in the example below:

```
>>> from scipy.stats import *
>>> # norm.cdf(t, mu, sigma) gives P(Y <= t) when Y has normal distribution
... # with mean mu and **standard deviation** sigma
...
>>> norm.cdf(0, 0, 1)
0.5
>>> norm.cdf(-1, 0, 1)
0.15865525393145707
>>> norm.cdf(2, 0, 1)
0.9772498680518208
>>>
```