STAT 400 : DISCUSSION SECTION BD3 AND BD4 HANDOUT 3

SUBHADEEP PAUL

1. Theory and Discussions

1.1. Continuous random variables: Concepts to remember.

1.1.1. *Difference with discrete RVs :*

- Sample space is continuous instead of discrete. Recall : we spoke about extending the definition of probability from numbers to intervals and defined probability as either length or area. So probability of an interval (a,b) is defined as the area under the graph of f(x) in between the two points a and b, i.e $p(a < x < b) = \int_a^b f(x)$. Probability of a single point x is 0.
- PMF(X=x) meant probability of X=x, but pdf(x) doesn't mean probability ! so pdf at any point x can be greater than 1 as well(but always non negative).

1.1.2. *Similarities with discrete RVs :*

- Probability of the entire sample space is 1. so the area under the entire curve $\int_{-\infty}^{\infty} f(x)dx = 1$. Definition of expectation and variance is same with $\sum$ replaced by $\int$ over the entire sample space. So, $E(X) = \int_S xf(x)dx$. CDF is defined in the same way as the discrete case, $P(X \leq x) = F_X(x) = \int_{-\infty}^{x} f(x)dx$.
- Take a look at Ex_3_3_2.pdf and continuous.pdf in compass 2g.

1.2. Median, percentiles, quantiles and Mode.

- The $100k$ th percentile is $F_X(\pi_k) = P(X \leq \pi_k) = k$, e.g $80th$ percentile is $F_X(\pi_{0.80}) = P(X \leq \pi_{0.80}) = 0.80$. The 3 quantiles are 25th, $50th$ and $75th$ percentiles. Median is the $50th$ percentile. So Median(X), π_{median} is given by $F_X(\pi_{median}) = P(X \leq \pi_{median}) = 0.50$.
- Mode corresponds to the value of x in a distibution which has highest frequency. For discrete distributions is defined as the point with highest probability and for continuous distribution it is the point with largest values of f(x), i.e it is a maxima of f(x). [Recall from math : to find maxima of a function you need to differentiate it and equate to 0]

1.3. Moment generating function:

- Definition: $E(e^{tX})$, i.e $\sum e^{tx}p(x)$ or $\int e^{tx}f(x)dx$. Note that this will be a function of t. Now differentiate and evaluate at $t = 0$. This is $E(X)$. Further differentiation gives higher order moments, $E(X^2)$ and so on.

1.4. Normal distribution:

- $X \sim N(\mu, \sigma^2)$, $\mu$is the mean, median, and mode. σ is the standard deviation.
- The great thing about normal distribution is if $X \sim N$, so does all its linear transformations, $Y = aX + b$. Clearly, $E(Y) = aE(X) + b$ and $Var(Y) = a^2Var(X)$. So we need to have the table for the simplest case $N(0, 1)$. So, if $X \sim N(\mu, \sigma^2)$, $z = \frac{X-\mu}{\sigma} \sim N(0, 1)$. (verify ! we did this in class and had a homework on this). Z

is called standard normal random variable and its pdf and cdf are denoted by $\phi(x)$, and, $\Phi(x)$ respectively. If you add two independent normal RVs, both the mean and the variance (NOT the sd) add up to give the mean and variance of the resulting RV.
- $\Phi(-x) = 1 - \Phi(x)$. $\Phi(0) = 0.5$. How to look up the Standard normal table? (discuss in class).

1.5. **Review for midterm:**
- Independence: What does A and B independent mean? $P(A|B) = P(A)$, which implies, $P(A \cap B) = P(A)P(B)$
- Go through Bayes' theorem, last week's handout on identifying distributions.

2. Problems

Problem 1. Suppose a random variable X has the following probability density function:

$$f(x) = \begin{cases} C|x-2| & 0 \le x \le 4 \\ 0 & otherwise \end{cases}$$

(1) What must the value of C be so that $f(x)$ is a probability density function?
(2) Find the cumulative distribution function, $P(X \le x) = F_X(x)$.
(3) Find the median of the probability distribution of X.
(4) Find mean, $E(X)$.
(5) Find the MGF, $M_X(t)$.

Note. The functional form of the pdf changes within the domain of x.

Problem 2. Now a variation of the earlier problem

$$f(x) = \begin{cases} \frac{2}{5}|x-2| & 0 \le x \le k \\ 0 & otherwise \end{cases}$$

(1) What must the value of k be so that $f(x)$ is a probability density function?
(2) Find the cumulative distribution function, $P(X \le x) = F_X(x)$. Or, find $P(X \le 2)$ and $P(x \le 4)$.
(3) Find the median of the probability distribution of X.
(4) Find mean, $E(X)$.

Problem 3. Suppose a discrete random variable X has the following probability distribution:

$$P(X = k) = \frac{(ln2)^k}{k!}, k = 1, 2, 3...$$

(1) Find the moment generating function of X, $M_X(t)$.
(2) Use this function to find $E(X)$ and Var(X).

Problem 4. A candy maker produces mints that have a label weight of 20.4 grams. Assume that thedistribution of the weights of these mints is N(21.37,0.16).

(1) Let X denote the weight of a single mint selected at random from the production line. Find P(X > 22.07).
(2) Suppose that 15 mints are selected independently and weighed. Let Y equal the number of these mints that weigh less than 20.857 grams. Find P($Y \le 2$).