Functions of Random Variables/Expectation and Variance

ECE 313

Probability with Engineering Applications

Lecture 25 - November 19, 1999

Professor Ravi K. Iyer

University of Illinois

Mean, Median, and Mode

- The distribution function *F(x)* or the density *f(x)* (or pmf $p(x_i)$) completely characterizes the behavior of a random variable *X*.
- Often, we need a more concise description such as a single number or a few numbers, instead of an entire function.
- Quantities most often used to describe a random variable *X* are
  - the **expectation** or the **mean**, *E[X].*
  - the **median**, any number *x* such that $P(X < x) \leq 1/2$ and $P(X > x) \geq 1/2$ and
  - the **mode**, any number *x* for which f*(x)* or $p(x_i)$ attains its maximum.
- The mean, median, and mode are often called **measures of central tendency** of a random variable *X*.

Expectation

- The expectation, *E[X]*, of a random variable *X* is defined by:

$$E[X] = \begin{cases} \sum_i x_i p(x_i) & \text{if } X \text{ is discrete} \\ \int_{-\infty}^{\infty} x f(x) dx & \text{if } X \text{ is continuous} \end{cases}$$

provided that the relative sum or integral is absolutely convergent;

$$\sum | x_1 | \, p(x_i) < \infty \quad \text{and} \quad \int_{-\infty}^{\infty} | x | \, f(x) dx < \infty$$

If the right hand side in is not absolutely convergent, then *E[X]* does not exist.

Expectation - Example 1

- Consider the problem of searching for a specific name in a table of names.
- A simple method is to scan the table sequentially until the name appears or is found missing.
- A program segment for sequential search:

```
var T :  array [0..n] of NAME;
Z :  NAME;
I :  0..n;
begin
   T[0] := Z;  {T[0] is used as a sentinel or marker}
   I := n;
   while Z ≠ T[I] do
     I := I - 1;
   if I > 0 then {found;  I points to Z}
    else{not found}.
end
```

Expectation - Example 1 (cont.)

- To analyze the time required for sequential search, let X be the discrete random variable denoting the number of comparisons "Z ≠ T[I]" made.
- The set of all possible values of X is {1, 2, …, n+1} and $X = n+1$ for unsuccessful searches.
- We consider a random variable Y that denotes the number of comparisons on a successful search (note that the value of X is fixed for unsuccessful searches).
- The set of all possible values of Y is {1, 2, …, n}.

Expectation - Example 1 (cont.)

- To compute the average search time for a successful search, we must specify the pmf of Y.
- In the absence of any specific information, it is natural to assume that Y is uniform over its range:

$$p_Y(i) = \frac{1}{n}, \qquad 1 \le i \le n.$$

then

$$E[Y] = \sum_{i=1}^{n} i p_Y(i) = \frac{1}{n}\frac{n(n+1)}{2} = \frac{n+1}{2}$$

- Thus, on the average, approximately half the table needs to be searched.

Expectation - Example 2

- The assumption of uniform distribution, used in the Example 1, rarely holds in practice.
- It is possible to collect statistics on access patterns and use empirical distributions to reorganize the table to reduce the average search time.
- Assume that the table search starts from the front.
- If α_i denotes the access probability for name *T*[*i*], then the average successful search time $E[Y] = \Sigma\, i\alpha_i$.
- $E[Y]$ is minimized when names in the table are in the order of nonincreasing access probabilities:

$$\alpha_1 \geq \alpha_2 \geq \ldots \geq \alpha_n$$

Expectation - Example 2 (cont.)

- In practice, for example, many tables follow Zipf's law:

$$\alpha_i = \frac{c}{i}, \qquad 1 \leq i \leq n,$$

where the constant *c* is determined from the normalization requirement:

$$\sum_{i-1}^{n} \alpha_i = 1$$

$$\text{Thus :} \quad c = \frac{1}{\sum_{i=1}^{n} \frac{1}{i}} = \frac{1}{H_n} \approx \frac{1}{1n(n)}$$

$$\text{where } H_n = \sum_{i=1}^{n} \frac{1}{i}.$$

Expectation - Example 2 (cont.)

- If the names in the table are ordered as shown on the previous slide, the average search time is:

$$E[Y] = \sum_{i=1}^{n} i\alpha_i = \frac{1}{H_n}\sum_{i=1}^{n} 1 = \frac{n}{H_n} \approx \frac{n}{1n(n)}$$

which is considerably less than the previous value (*n* + 1)/2, for large *n*.

Expectation - Example 3 (cont.)

- Thus, for example, if a component obeys an exponential failure law with parameter λ (failure rate) then its expected life, or its mean time to failure (MTTF), is $1/\lambda$.

- Similarly, if the interarrival times of jobs to a computer center are exponentially distributed with parameter λ (arrival rate), then the mean interarrival time is $1/\lambda$.

Moments

- Let X be a random variable and let define another random variable Y such that $Y=\phi(X)$.
- We want to compute E[Y], we could compute the pmf or pdf of Y using methods discussed earlier.
- It is easier to use the following expressions

$$E[Y]=E[\phi(X)]=\begin{cases}\sum_i \phi(x_i)p_x(x_i), & \text{if X is discrete,}\\ \int_{-\infty}^{\infty}\phi(x)f_x(x)dx, & \text{if X is continuous}\end{cases}$$

provided the sum or integral on the right-hand side is absolutely convergent:

Moments (cont.)

- A special case of interest is the power function $\phi(X)=X^k$, for *k=1,2,3,…,* $E[X^k]$ is known as the k^{th} *moment* of the random variable X.
- *Note that the first moment, E[X], is the ordinary expectation or the mean of X.*
- We can show that if X and Y are random variables that have matching corresponding moments of all orders (i.e., $E[X^k]=E[Y^k]$ for k=1,2,…), then X and Y have the same distribution.

Variance of a Random Variable

- To center the origin of measurement, it is convenient to work with powers of X-E[X].
  - We define the k^{th} central moment, μ_k, of the random variable X

by $\mu_k = E[(X - E[X])^k]$,

- Of special interest is the quantity:

$$\mu_2 = E\left[(X - E[X])^2\right]$$

known as the *variance* of X, Var[X], often denoted by σ^2.