13.42 Design Principles for Ocean Vehicles

Prof. A.H. Techet

Spring 2005

1. Random Variables

A random variable is a variable, x, whose value is assigned through a rule and a random experiment, ζ, that assigns a priori a value to the outcome of each experiment, $A_1, A_2, A_3, \ldots$ This rule states that

$$x(A_1) = x_1$$
$$x(A_2) = x_2$$
$$\cdots$$
$$x(A_n) = x_n$$

One example of a random variable is a Bernoulli random variable which assigns either a 1 or 0 to the outcome. For example, toss a "fair" coin. If it lands heads up you get one dollar, if it land tails up you loose a dollar. The amount won or lost in this case is the random variable.

Symbolically, $x(\zeta)$ denotes the random variable which is a "function" of the random event $\zeta = \{A_1, A_2, \ldots, A_n\}$ which has associated probabilities: $p(A_1) = p_1$, $p(A_2) = p_2$, etc.

$$A_1 \longrightarrow x_1, p_1$$
$$A_2 \longrightarrow x_2, p_2$$
$$\vdots$$
$$A_n \longrightarrow x_n, p_n$$

The variables x_i are the values of the random variable, A_i, the possible events in the event space, and p_i is the probability of event A_i.

EXPECTED VALUE of the random variable can be thought of as follows: after many (M) repetitions of a random experiment in which event A_1 occurred d_1 times, A_2 occurred d_2 times, and so on to A_n occurred d_n times, the total number of experiments is simply

$$M = d_1 + d_2 + d_3 + \cdots + d_n. \tag{19}$$

If a weight, or cost, x_i, is assigned to each event, A_i, then the total cost of all of the events is

$$x_T = d_1 x_1 + d_2 x_2 + \cdots + d_n x_n. \tag{20}$$

Then given p_i, the probability of event A_i, the expected value of the event is

$$\bar{x} = E\{X(\zeta)\} = \sum_{i=1}^{N} p_i x_i. \tag{21}$$

Hence the **AVERAGE INCOME** per trial is

$$\bar{x} = \frac{x_T}{M} \tag{22}$$

As $m \to \infty$: $\frac{d_i}{M} \approx p_i$. In other words the number of occurrences of each event, d_i, divided by the total number of events, M, is equal to the probability of the event, p_i, and the expected value of x is the average income as defined in 22 as $M \to \infty$

$$\bar{x} = x_1 p_1 + x_2 p_2 + \cdots + x_n p_n . \tag{23}$$

Expected Value Properties

$$E\{x + y\} = E\{x\} + E\{y\}$$

$$E\{C\} = C \text{ ; C is a constant}$$

$$E\left\{g(x(\zeta))\right\} = \sum_{i=1}^{n} g(x_i)\ p_i$$

Properties of Variance

$$\begin{aligned} Vx \quad &= E\left\{\left[x(\zeta) - \bar{x}\right]^2\right\} \\ &= E\left\{x^2(\zeta) - 2x(\zeta)\bar{x} + \bar{x}^2\right\} \\ &= E\{x(\zeta)^2\} - 2\bar{x}E\{x(\zeta)\} + \bar{x}^2 \\ &= E\left\{x^2(\zeta)\right\} - \bar{x}^2 \end{aligned}$$

The Standard deviation is defined as the square root of the variance: $\sigma = \sqrt{V}$.

2. Probability Distribution

Discrete Random Variable: possible values are a finite set of numbers or a countable set of numbers.

Continuous Random Variable: possible values are the entire set or an interval of the real numbers. The probability of an outcome being any *specific point* is zero (improbable but not impossible).

EXAMPLE: On the first day of school we observe students at the campus bookstore buying computers. The random variable x is zero if a desktop is bought or one if the laptop is bought. If 20% of all buyers purchase laptops then the pmf of x is

$$p(0) = p(X = 0) = \text{ p(next customer buys a desktop) } = 0.8$$
$$p(1) = p(X = 1) = \text{ p(next customer buys a laptop) } = 0.2$$
$$p(x) = p(X = x) = 0 \text{ for } x \neq 0 \text{ or } 1.$$

Probability Density Function (pdf): of a continuous random variable, x, is defined as the probability that x takes a value between x_o and $x_o + dx$, divided by dx, or

$$f_x(x_o) = p(x_o \leq x < x_o + dx)/dx. \qquad (24)$$

This must satisfy $f_x(x) \geq 0$ for all x where $\int_{-\infty}^{\infty} f_x(x) = 1$ is the area under the entire graph $f_x(x)$. It should be noted that a PDF is NOT a probability.

3. Cumulative Distribution Function (CDF)

At some fixed value of x we want the probability that the observed value of x is at *most* x_o. This can be found using the cumulative distribution function, $P(x)$.

Discrete Variables: The cumulative probability of a discrete random variable x_n with probability $p(x)$ is defined for all x as

$$F(x \leq x_k) = \sum_{j=1}^{k} p(x_k) \qquad (25)$$

Continuous Variables: The CDF, $F(x)$, of a continuous random variable X with pdf $f_x(x)$ is defined for all x as

$$F_x(x_o) = p(X \leq x_o) = \int_{-\infty}^{x_o} f_x(y)dy \qquad (26)$$

which is the area under the probability density curve to the left of value x_o. Note that $F(x_o)$ is a probability in contrast to the PDF. Also

$$F(x_o) = p(x \le x_o) = \int_{-\infty}^{x_o} f_x(x)dx \tag{27}$$

and

$$f_x(x_o) = \frac{dF(x_o)}{dx}. \tag{28}$$

Let x be a continuous random variable with a pdf $f_x(x)$ and cdf $F(x)$ then for any value, a,

$$p(x > a) = 1 - F(a); \tag{29}$$

and for any two numbers, a and b,

$$p(a \le x \le b) = F(b) - F(a) \tag{30}$$

Expected Value: The expected value of a continuous random variable, *x*, with pdf $f_x(x)$ is

$$\mu_x = E(x) = \int_{-\infty}^{\infty} x \; f_x(x)dx \tag{31}$$

If x is a continuous random variable with pdf $f_x(x)$ and $h(x)$ is any function of that random variable then

$$E[h(x)] = \mu_{h(x)} = \int_{-\infty}^{\infty} h(x) \; f_x(x)dx \tag{32}$$

Conditional Expectations: The expected value of the random variable given that the variable is greater than some value.

Example:

Variance: The variance of a continuous random variable, x, with pdf $f_x(x)$ is

$$\sigma_x^2 = V\{x\} = E\{[x-\bar{x}]^2\} = \int_{-\infty}^{\infty} (x-\mu_x)^2 \; f_x(x)dx \tag{33}$$

4. Functions of Random Variables

Given a random variable, $X(\zeta)$ or pdf, $f_x(x)$, and a function, $y = g(x)$, we want to find the probability of some y, or the pdf of y, $f_y(y)$.

$$F(X \le x_o) = F(y(x) \le g(x_o)) \tag{34}$$

The probability that the random variable, X, is less than some value, x_o, is the same as the probability that the function $y(x)$ is less than the at function evaluated at x_o.

EXAMPLE: Given $y = \alpha x + b$ and the pdf $f_x(x)$ for all $\alpha > 0$, then $\alpha x + b < y_o$ for $x \le \frac{y_o - b}{\alpha}$ and

$$F(y \le y_o) = \int_{-\infty}^{\frac{y_o-b}{\alpha}} \; f_x(x)dx \tag{35}$$

EXAMPLE: Given $y = x^3$: $F(X \le x_o) = F(y \le x_o^3)$.

If $y \to x$ has one solution and pdfs f_y and f_x

$$f_y \, |dy| = f_x \, |dx| \tag{36}$$

$$f_y = f_x / \frac{|dy|}{|dx|} \tag{37}$$

If $y \to x_1, x_2, \cdots, x_n$ then

$$f_y = \frac{f_x(x_1)}{\left|\frac{dy(x_1)}{dx}\right|} + \cdots + \frac{f_x(x_n)}{\left|\frac{dy(x_n)}{dx}\right|} \tag{38}$$

5. Central Limit Theorem

Let $x_1, x_2, \ldots, x_n$ be random samples from an arbitrary distribution with mean, μ, and variance, σ^2. If *n* is sufficiently large, $\bar{x}$ has an approximately *normal* distribution. So as $n \to \infty$, if $f_x(x)$ can be approximated by a Gaussian distribution. then

$$\bar{x} = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} x_i \quad (39$$

and

$$\sigma^2(x) = \frac{1}{n} \sum_{i=1}^{n} \sigma_i^2 \tag{40}$$