Lecture 2 - Objective and Outcomes

Random variables allow us to associate real numbers with events. This week we discuss the implications of this mapping and introduce the idea of a distribution.

- random variables
- cumulative distribution functions
- discrete variables and mass functions

- continuous variables and density functions
- expectation and variance
- expectation operator

After reviewing the notes you should:

- understand the distinction between a random variable and an instance of a random variable
- be able to derive mass/density functions from distribution functions and vice versa
- be able to derive mean and variance for a large number of distributions

Properties of distribution functions

1. Limits:
   $\lim_{x \to -\infty} F(x) = 0$ and $\lim_{x \to \infty} F(x) = 1$.

2. Non-decreasing:
   if $x < y$ then $F(x) \leq F(y)$.

3. Right continuous:
   $\lim_{h \downarrow 0} F(x + h) = F(x)$.

Probabilities from distribution functions

1. $P(X > x) = 1 - F(x)$.

2. $P(x < X \leq y) = F(y) - F(x)$.

3. $P(X < x) = \lim_{h \downarrow 0} F(x - h) = F(x-)$.

4. $P(X = x) = F(x) - F(x-)$.

Properties of mass functions

1. $f(x) = F(x) - F(x-)$.

2. $\sum_{i:x_i \leq x} f(x_i) = F(x)$.

3. $\sum_i f(x_i) = 1$.

4. $f(x) = 0$ if $x \notin \{x_1, x_2, \ldots\}$.

Properties of continuous random variables

1. $P(X = x) = 0$ for all $x \in \mathbb{R}$.

2. $\int_{-\infty}^{\infty} f(x)dx = 1$.

3. $\int_{a}^{b} f(x)dx = P(a < X \leq b)$.

Unified notation

$$P(a < X \leq b) = \int_a^b dF(x)$$
$$= \begin{cases} \sum_{i:a<x_i\leq b} f(x_i), & \text{if } X \text{ discrete,} \\ \int_a^b f(x)dx, & \text{if } X \text{ continuous.} \end{cases}$$

Properties of expectation

For constant a and random variables X and Y.

1. $E(a) = a$.

2. $E(aX) = aE(X)$.

3. $E(X + Y) = E(X) + E(Y)$.

4. If $X \geq Y$ then $E(X) \geq E(Y)$.

Inequalities involving expectation

1. Markov inequality:
   random variable $Y \geq 0$, constant $a > 0$;

$$P(Y \geq a) \leq E(Y)/a.$$

2. Chebyshev inequality: constant $a > 0$;

$$P(|X - E(X)| \geq a) \leq \sigma^2/a^2.$$

Moments

If r is a positive integer then the r^{th} *moment*, m_r, of X is

$$m_r = E(X^r).$$

The r^{th} *central moment*, μ_r is

$$\mu_r = E((X - m_1)^r).$$

Properties of moments

1. $m_1 = E(X) = \mu$ (mean); $\mu_1 = 0$.
2. $\mu_2 = E(X^2) - E(X)^2 = \mathrm{var}(X) = \sigma^2$ (variance).
3. Coefficient of skewness: $\gamma_1 = E[(X-\mu)^3]/\sigma^3 = \mu_3/\mu_2^{\frac{3}{2}}$.
4. Coefficient of kurtosis:
   $\gamma_2 = (E[(X-\mu)^4]/\sigma^4) - 3 = (\mu_4/\mu_2^2) - 3$.

Moment generating function

The MGF of a random variable X is a function $M : \mathbb{R} \longrightarrow [0, \infty)$ given by

$$M(t) = E(e^{tX}),$$

where $M(t) < \infty$ for $|t| < h$ and some $h > 0$.

1. Evaluation via integration:

$$M(t) = \int_{-\infty}^{\infty} e^{tx} dF(x) = \begin{cases} \sum_i e^{tx_i} f(x_i), & \text{if } X \text{ discrete,} \\ \int_{-\infty}^{\infty} e^{tx} f(x) dx, & \text{if } X \text{ continuous.} \end{cases}$$

2. Taylor expansion:

$$M(t) = 1 + tE(X) + \frac{t^2}{2!}E(X^2) + \ldots + \frac{t^r}{r!}E(X^r) + \ldots = \sum_{j=0}^{\infty} \frac{E(X^j)}{j!} t^j.$$

3. Derivatives at zero:

$$M^{(r)}(0) = \frac{d^r}{dt^r}M(t)\bigg|_{t=0} = E(X^r) = m_r.$$

4. Uniqueness: If $M_X(t) = M_Y(t)$ for all $|t| < h$ and some $h > 0$, then $F_X(x) = F_Y(x)$ for all $x \in \mathbb{R}$.

Cumulant generating function

: of a random variable X with moment generating function $M(t)$, is defined as

$$K(t) = \log M(t),$$

The r^{th} *cumulant*, κ_r, is the coefficient of $t^r/r!$ in the expansion of the cumulant generating function $K(t)$;

$$\kappa_r = K^{(r)}(0) = \frac{d^r}{dt^r} K(t)\bigg|_{t=0}.$$

1. $\kappa_1 = m_1 = \mu$ (mean, first moment).

2. $\kappa_2 = \mu_2 = \sigma^2$ (variance, second central moment).

3. $\kappa_3 = \mu_3$ (third central moment).

4. $\kappa_4 + 3\kappa_2^2 = \mu_4$ (fourth central moment).

Probability generating function

: of a discrete random variable X, taking non-negative integer values, is defined as

$$G(t) = E(t^X) = \sum_{i=0}^{\infty} t^i P(X = i).$$

1. $P(X = 0) = G(0)$.

2. $E(X) = G'(1)$.

3. $E(X(X-1)\ldots(X-k+1)) = G^{(k)}(1)$.

Characteristic function

: of a random variable X is the function $\phi : \mathbb{R} \longrightarrow \mathbb{C}$ defined by

$$\phi(t) = E(e^{itX})$$

where $i = \sqrt{-1}$. Note that $E(X^k) = i^{-k}\phi^{(k)}(0)$.

Functions of random variables

If X is a random variable with distribution function F_X and g is a real-valued function, then $Y = g(X)$ is a random variable.

- If g is strictly increasing $F_Y(y) = F_X(g^{-1}(y))$.
- If g is strictly decreasing $F_Y(y) = 1 - F_X(g^{-1}(y))$.