Tutorial 9 (March 15^{th} | 2017)

Concept Review

1) A random variable is represented by capital letter and takes values randomly according to some distribution. Example: X, Y, Z.

2) The random variable (RV) may be discrete (i.e. taking finitely or countably many values) or continuous (i.e. taking uncountably many values).
Example: $X \sim \text{Binomial}(n, p)$ or $Y \sim \text{Uniform}(\theta_1, \theta_2)$

3) Probability is represented either by a probability mass function (pmf) for a discrete RV or an integral of the probability density function (pdf) for a continuous RV.
Example: If $X \sim \text{Binomial}(n,p)$ then X takes values $0, 1, \ldots, n$ where $\mathbb{P}(X = x) = p(x) = \binom{n}{x} p^x (1-p)^{n-x}$ where x is some numbers $0, 1, 2, \ldots, n$. If $X \sim \text{Uniform}(5, 6)$ then
$$\mathbb{P}(5.5 \le X \le 5.6) = \int_{5.5}^{5.6} f(x)\,dx = \int_{5.5}^{5.6} \frac{1}{6-5}\,dx = \int_{5.5}^{5.6} 1\,dx = 0.1$$

4) Expectation is a measure of "center" and is a sum for discrete RV or an integral for continuous RV.
Example: $X \sim \text{Binomial}(n,p)$, $\mathbb{E}(X) = \sum_{x=0}^{n} x p(x) = \sum_{x=0}^{n} x \binom{n}{x} p^x (1-p)^{n-x}$.
if $Y \sim \text{Uniform}(5,6)$, $\mathbb{E}(Y) = \int_5^6 y f(y)\,dy = \int_5^6 y\,dy$.

5) The Cumulative Distribution function (cdf) is a function over the support of a RV. It is given as $F(x) = \mathbb{P}(X \le x)$ where again X is the RV and x is a number. For discrete RV, we add up all the pmf for which $X \le x$ and for a continuous RV, we integrate the pdf from $-\infty$ to x.
Example: $X \sim \text{Binomial}(5, 0.2)$, $F(3.3) = \mathbb{P}(X \le 3.3)$
$$= p(0) + p(1) + p(2) + p(3)$$
$Y \sim \text{Normal}(\mu, \sigma^2)$, $F(3.2) = \mathbb{P}(Y \le 3.2) = \int_{-\infty}^{3.2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-\mu)^2}{2\sigma^2}}\,dy$

6) Functions of RV are still RV. If X is a RV and g is a function then $Y = g(X)$ is also a RV. All operations on the previous page are analogous for $g(X)$.

$$E(Y) = E(g(X)) = \begin{cases} \sum_{supp(X)} g(x)\,p(x) & , X \text{ discrete} \\ \int_{-\infty}^{\infty} g(x)\,f(x)\,dx & , X \text{ continuous} \end{cases}$$

* If g is known to be strictly increasing then we can relate probability statements of $g(X)$ back to X

Example: $\mathbb{P}(g(X) \leq x) = \mathbb{P}(X \leq g^{-1}(x))$
$= F(g^{-1}(x))$ where again, x is a number and X is a RV.

$$\mathbb{P}(a \leq g(X) \leq b) = \mathbb{P}(g^{-1}(a) \leq X \leq g^{-1}(b))$$

and if X is discrete, we add up all the pmf values between $g^{-1}(a)$ and $g^{-1}(b)$ and if X is continuous, we integrate the pdf between $g^{-1}(a)$ and $g^{-1}(b)$, i.e. $\int_{g^{-1}(a)}^{g^{-1}(b)} f(x)\,dx$

7) Equivalently Distributed Random variables are those in which the CDFs are equal. That is, if X and Y have the same distribution then $\mathbb{P}(X \leq x) = \mathbb{P}(Y \leq x)$ for all values x.

8) Proving a statement is true for all integer $n \geq 1$ can either be done directly or by using a technique known as induction. Suppose we are required to prove $1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}$.

Proving the statement directly for all $n \geq 1$ is difficult. We will prove by induction.

Proofs by induction are divided into 3 steps.

1. Base Case (prove the statement for the simplest case)
2. Induction Hypothesis (assume the statement is true for some integer $n = k \geq 1$)
3. Induction Step (prove the statement for $n = k+1$ using the induction hypothesis)

We will prove $1 + 2 + 3 + \dots + n = \frac{n(n+1)}{2}$ by induction, $n \geq 1$.

1. Base Case: $n = 1$.

   Left Hand Side (LHS) = 1

   Right Hand Side (RHS) = 1. The statement is true for $n = 1$.

2. Induction Hypothesis: Suppose the statement is true for $n = k$.

   That is, $1 + 2 + \dots + k = \frac{k(k+1)}{2}$.

3. Induction Step: We want to show $1 + 2 + \dots + k + 1 = \frac{(k+1)(k+2)}{2}$

We start from the LHS:

$$1 + 2 + \dots + k + 1 = \underbrace{1 + 2 + \dots + k}_{} + k + 1$$

Look familiar? We assumed in our induction hypothesis that this was equal to $\frac{k(k+1)}{2}$.

Thus, the LHS reduces to the following:

$$\underbrace{1 + 2 + \dots + k}_{= \frac{k(k+1)}{2}} + k + 1 = \frac{k(k+1)}{2} + k + 1 = \frac{k(k+1) + 2k + 2}{2}$$

$$= \frac{k^2 + 3k + 2}{2}$$

$$= \frac{(k+1)(k+2)}{2} = \text{RHS}$$

Therefore, we have proven that the LHS = RHS for all $n \geq 1$.