

## 01. Basic Concepts of Probability

### Event Operations

- **Mutually Exclusive** -  $A \cap B = \emptyset$
- **Contained** -  $A \subset B$
- **Equivalence** -  $A \subset B$  and  $A \supset B \rightarrow A = B$
- **Distributive** -  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- **DeMorgan's** -  $(A \cup B)' = A' \cap B'$
- $A = (A \cap B) \cup (A \cap B')$

### Counting Methods

- **Multiplication Principle** - Given r experiments performed sequentially and each has  $n_1, n_2, \dots, n_r$  outcomes. After r experiments, there are  $n_1 n_2 \dots n_r$  outcomes.
- **Addition Principle** - Given experiment can be done in k different ways and each has  $n_1, n_2, \dots, n_r$  ways. There are  $n_1 + n_2 + \dots + n_k$  total ways.
- **Permutation** -  ${}_nP_r = \frac{n!}{(n-r)!}$
- **Combination** -  $\binom{n}{r} = \frac{n!}{(n-r)!r!}$

### Probability

#### Axioms of Probability

1. For any event A,  $0 \leq P(A) \leq 1$
  2.  $P(S) = 1$
  3. If  $A \cap B = \emptyset$ , then  $P(A \cup B) = P(A) + P(B)$
- $P(A') = 1 - P(A)$
  - $P(A) = P(A \cap B) + P(A \cap B')$
  - $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
  - If  $A \subset B$ , then  $P(A) < P(B)$

### Finite Sample Space with Equally Likely Outcomes

Given sample space  $S = \{a_1, \dots, a_k\}$  and all outcomes are equally likely, i.e.  $P(a_1) = \dots = P(a_k)$ :

$$\text{For any event } A \subset S, P(A) = \frac{\text{No. of sample points in } A}{\text{No. of sample points in } S}$$

### Conditional Probability

$$P(B|A) = \frac{P(B \cap A)}{P(A)} = \frac{P(A|B)P(B)}{P(A)}$$

### Independence

- $A \perp B \leftrightarrow P(A \cap B) = P(A)P(B)$
- $A \perp B \leftrightarrow P(A|B) = P(A)$

### Law of Total Probability

- **Partition** - If  $A_1, \dots, A_n$  are mutually exclusive events and  $\bigcup_{i=1}^n A_i = S$ , then  $A_1, \dots, A_n$  are partitions
- If  $A_1, \dots, A_n$  are partitions of S, then for any event B:

$$P(B) = \sum_{i=1}^n P(B \cap A_i) = \sum_{i=1}^n P(B|A_i)P(A_i)$$

### Bayes' Theorem

Let  $A_1, \dots, A_n$  be partitions of S. For any event B:

$$P(A_k|B) = \frac{P(B|A_k)P(A_k)}{\sum_{i=1}^n P(B|A_i)P(A_i)}$$

## 02. Random Variables

- Motivation: Assign value to outcome of experiment
- **Random Variable** - Let S be sample space. Function X which maps  $\mathbb{R}$  to every  $s \in S$

### Probability Distribution

- Probability assigned to each possible X
- Given RV X with range of  $R_x$ :
  - **Discrete** - Numbers in  $R_x$  are finite or countable
  - **Continuous** -  $R_x$  is interval

### Discrete Probability Distribution

- **Probability Function** - Given  $R_x = \{x_1, \dots\}$ . For each  $x_i$ , there's some probability that  $X = x_i$ :

$$f(x) = P(X = x)$$

- *p.f.* must satisfy:
  1.  $f(x_i) = P(X = x_i)$  for  $x_i \in R_x$
  2.  $f(x_i) = 0$  for  $x_i \notin R_x$
  3.  $\sum_{i=1}^{\infty} f(x_i) = 1$
  4.  $\forall B \subseteq \mathbb{R}, P(X \in B) = \sum_{x_i \in B \cap R_x} f(x_i)$
- **Probability Distribution** - Collection of pairs  $(x_i, f(x_i))$

### Continuous Probability Distribution

- **Probability Function** - Given  $R_x$  is interval. Quantifies probability that X is in some range.
- *p.f.* must satisfy:
  1.  $f(x) \geq 0$
  2.  $f(x) = 0$  for  $x \notin R_x$
  3.  $\int_{R_x} f(x)dx = 1$
  4.  $\forall a, b \text{ s.t. } a \leq b, P(a \leq X \leq b) = \int_a^b f(x)dx$
- Note:  $P(X = x_0) = \int_{x_0}^{x_0} f(x)dx = 0$

### Cumulative Distributive Function

Given RV X, which can be discrete or continuous:

$$F(x) = P(X \leq x)$$

- $F(x)$  is non-decreasing and  $0 \leq F(x) \leq 1$
- **For discrete RV**: Step function

$$F(x) = \sum_{t \in R_x; t \leq x} f(t)$$

- $P(a \leq X \leq b) = F(b) - \lim_{x \rightarrow a^-} F(x)$
- $0 \leq f(x) \leq 1$

- **For continuous RV**:

$$F(x) = \int_{-\infty}^x f(t)dt$$

$$f(x) = \frac{d(F(x))}{dx}$$

- $P(a \leq X \leq b) = P(a < X < b) = F(b) - F(a)$
- $0 \leq f(x)$  e.g.  $f(x) = 3x^2$  is a valid *p.f.* since  $\int_{R_x} f(x)dx = 1$

### Expectation of Random Variable

- **Mean of discrete RV**:

$$\mu = E(X) = \sum_{x \in R_x} x_i f(x_i) = \sum_{i=1}^{\infty} P(X \geq i)$$

- Let  $g$  be some function.  $E(g(x)) = \sum_{x \in R_x} g(x)f(x)$

- **Mean of continuous RV**:

$$\mu = E(X) = \int_{x \in R_x} x f(x)dx$$

- Let  $g$  be some function.  $E(g(x)) = \int_{x \in R_x} g(x)f(x)dx$

- $E(aX + b) = aE(X) + b$

- Linearity of expectation:  $E(X + Y) = E(X) + E(Y)$

### Variance of Random Variable

$$\sigma_X^2 = V(X) = E((X - \mu_X)^2)$$

- **Variance of discrete RV**:

$$V(X) = \sum_{x \in R_x} (x - \mu_X)^2 f(x)$$

- **Variance of continuous RV**:

$$V(X) = \int_{x \in R_x} (x - \mu_X)^2 f(x)dx$$

- $V(X) = 0$  when X is a constant

- $V(aX + b) = a^2 V(X)$

- $V(X) = E(X^2) - (E(X))^2$

- **Standard Deviation** -  $\sigma_X = \sqrt{V(X)}$

## 03. Joint Distributions

- Motivation: What if interested in more than 1 RV simultaneously?

- Given sample space S. Let X and Y be functions mapping  $s \in S \rightarrow \mathbb{R}$ :

$(X, Y)$  is 2D random vector

Range space:  $R_{X,Y} = \{(x, y) | x = X(s), y = Y(s), s \in S\}$

- **Discrete 2D RV** - If no. of possible values of  $(X(s), Y(s))$  are finite or countable
- **Continuous 2D RV** - If no. of possible values of  $(X(s), Y(s))$  can be any value in Euclidean space  $\mathbb{R}^2$
- If both X and Y are discrete/continuous, then  $(X, Y)$  is discrete/countinuous.

Joint Probability Function

• For discrete:

f\_{X,Y}(x,y) = P(X = x, Y = y)

- f\_{X,Y}(x,y) ≥ 0 for any (x,y) ∈ R\_{X,Y}
- f\_{X,Y}(x,y) = 0 for any (x,y) ∉ R\_{X,Y}
- ∑\_{i=1}^∞ ∑\_{j=1}^∞ P(X = x\_i, Y = y\_i) = 1
- Let A ⊆ R\_{X,Y}. P((X,Y) ∈ A) = ∑ ∑\_{(x,y) ∈ A} f\_{X,Y}(x,y)

• For continuous:

P(a ≤ X ≤ b, c ≤ Y ≤ d) = ∫\_a^b ∫\_c^d f\_{X,Y}(x,y) dy dx

- f\_{X,Y}(x,y) ≥ 0 for any (x,y) ∈ R\_{X,Y}
- f\_{X,Y}(x,y) = 0 for any (x,y) ∉ R\_{X,Y}
- ∫\_{-∞}^∞ ∫\_{-∞}^∞ f\_{X,Y}(x,y) dx dy = 1

Marginal Probability Function

Let (X,Y) be a 2D RV with joint probability function f\_{X,Y}(x,y):

If Y is discrete, f\_X(x) = ∑\_y f\_{X,Y}(x,y)

If Y is continuous, f\_X(x) = ∫\_{-∞}^∞ f\_{X,Y}(x,y) dy

- f\_Y(y) defined similarly
- Intuition: Marginal distribution for X ignores presence of Y
- f\_X(x) is a p.f.

Conditional Distribution

Let (X,Y) be a 2D RV with joint probability function f\_{X,Y}(x,y). Then ∀x s.t. f\_X(x) > 0:

f\_{Y|X}(y|x) = f\_{X,Y}(x,y) / f\_X(x)

- Intuition: Distribution of Y given X = x
- Only defined for x s.t. f\_X(x) > 0
- f\_{Y|X}(y|x) is a p.f. if we fix x
- But, f\_{Y|X}(y|x) is not a p.f. for x
- P(Y ≤ y | X = x) = ∫\_{-∞}^y f\_{Y|X}(y|x) dy
- E(Y | X = x) = ∫\_{-∞}^∞ y f\_{Y|X}(y|x) dy

Independent Random Variables

X ⊥ Y ↔ ∀x,y, f\_{X,Y}(x,y) = f\_X(x) f\_Y(y)

- Necessary condition: R\_{X,Y} must be a product space. Else, dependent.

Properties

Suppose X,Y are independent RV:

- If A, B ⊆ ℝ, then events X ∈ A and Y ∈ B are independent:

P(X ∈ A; Y ∈ B) = P(X ∈ A) P(Y ∈ B)

- g\_1(X) and g\_2(Y) are independent
- Independence is related with conditional distribution:

f\_X(x) > 0 → f\_{Y|X}(y|x) = f\_Y(y)

f\_Y(y) > 0 → f\_{X|Y}(x|y) = f\_X(x)

Quick way to check independence

1. R\_{X,Y} is a product space. i.e. R\_X does not depend on Y and vice versa.
2. f\_{X,Y}(x,y) can be written as c g\_1(x) g\_2(y) where g\_1 depends on x only and g\_2 depends on y only.
3. For discrete: f\_X(x) = ∑\_{t ∈ R\_X} g\_1(t) / g\_1(x)
4. For continuous: f\_X(x) = ∫\_{t ∈ R\_X} g\_1(t) dt / ∫\_{t ∈ R\_X} g\_1(t) dt

Expectation

Given 2 variable function g(x,y):

If (X,Y) is discrete, E(g(X,Y)) = ∑\_x ∑\_y g(x,y) f\_{X,Y}(x,y)

If (X,Y) is continuous, E(g(X,Y)) = ∫\_{-∞}^∞ ∫\_{-∞}^∞ g(x,y) f\_{X,Y}(x,y) dy dx

- E(XY) = E(X)E(Y) if X ⊥ Y

Covariance

cov(X,Y) = E((X - E(X))(Y - E(Y)))

If (X,Y) is discrete, cov(X,Y) = ∑\_x ∑\_y (x - μ\_X)(y - μ\_Y) f\_{X,Y}(x,y)

If (X,Y) is cont., cov(X,Y) = ∫\_{-∞}^∞ ∫\_{-∞}^∞ (x - μ\_X)(y - μ\_Y) f\_{X,Y}(x,y) dx dy

- cov(X,Y) = E(XY) - E(X)E(Y)
- X ⊥ Y → cov(X,Y) = 0. But converse is not always true.
- cov(aX + b, cY + d) = (ac)cov(X,Y)
- V(aX + bY) = a^2V(X) + b^2V(Y) + 2abcov(X,Y)
- X ⊥ Y → V(X + Y) = V(X) + V(Y)

04. Special Probability Distributions

Discrete Uniform Distribution

- If X has values x\_1, x\_2, ..., x\_k with equal probability
- p.f.: f\_X(x) = 1/k where x = x\_1, ..., x\_k and 0 otherwise
- Expectation: μ\_X = E(X) = ∑\_{i=1}^k x\_i f\_X(x\_i) = 1/k ∑\_{i=1}^k x\_i
- Variance: σ\_X^2 = V(X) = E(X^2) - (E(X))^2 = 1/k ∑\_{i=1}^k x\_i^2 - μ\_X^2

Bernoulli

- **Bernoulli Trial** - Random experiment with 2 possible outcomes (success and failure)

Bernoulli Random Variable

- Number of successes in Bernoulli trial (Either 1 or 0)
- Let 0 ≤ p ≤ 1 be the probability of success in Bernoulli trial

f\_X(x) = P(X = x) = { p      x = 1  
                         1 - p    x = 0  
                         0      otherwise

- f\_X(x) = p^x(1 - p)^{1-x} for x = 0 or 1
- Notation: X ~ Ber(p) and q = 1 - p
- μ\_X = E(X) = p and σ\_X^2 = V(X) = p(1 - p)

Bernoulli Process

- Sequence of repeatedly performed independent and identical Ber. trials
- Generates sequence of independent and identically distributed (i.i.d.) Ber. RVs: X\_1, X\_2, ...

Binomial Distribution

- **Binomial RV** - Counts the number of successes in n trials in a Ber. process
- Given n trials with each trial having probability p of success:

P(X = x) = (n choose x) p^x (1 - p)^{n-x}

- Notation: X ~ B(n,p)
- E(X) = np and V(X) = np(1 - p)

Negative Binomial Distribution

- Let X = Number of i.i.d. Bernoulli(p) trials until kth success occurs

P(X = x) = (x-1 choose k-1) p^k (1 - p)^{x-k}

- Notation: X ~ NB(k,p)
- E(X) = 1/p and V(X) = (1-p)/p^2

Geometric Distribution

- Let X = Number of i.i.d. Bernoulli(p) trials until 1st success occurs

P(X = x) = p(1 - p)^{x-1}

- Notation: X ~ G(p)
- E(X) = 1/p and V(X) = (1-p)/p^2

Poisson Distribution

- **Poisson RV** - Denotes number of events happening in fixed period of time or fixed region

P(X = k) = e^{-λ} λ^k / k!

- Notation: X ~ Poisson(λ) where λ > 0 is expected number of occurrences during given period/region
- E(X) = λ and V(X) = λ

Poisson Process

- Continuous time process, where we count number of occurrences within some interval of time
- Given Poisson process with rate parameter α:
  - Expected number of occurrences in interval of length T is αT
  - No simultaneous occurrences
  - Number of occurrences in disjoint intervals are independent
- Number of occurrences in any interval T of Poisson process follows Poisson(αT) distribution

Poisson Approximation of Binomial Distribution

Let X ~ B(n,p). Suppose n → ∞ and p → 0 s.t. λ = np remains constant. Then X ~ Poisson(λ) approximately.

lim\_{p→0; n→∞} P(X = x) = e^{-np} (np)^x / x!

- Approximation is good when n ≥ 20 and p ≤ 0.05, or n ≥ 100 and np ≤ 10

Continuous Uniform Distribution

X follows uniform distribution over interval  $(a, b)$  if  $p.f.$  is given by:

f\_X(x) = { 1/(b-a) if a <= x <= b, 0 otherwise }

• Notation:  $X \sim U(a, b)$

•  $E(X) = \frac{a+b}{2}$  and  $V(X) = \frac{(b-a)^2}{12}$

•  $c.d.f.$  is given by:

F\_X(x) = { 0 if x < a, (x-a)/(b-a) if a <= x <= b, 1 if x > b }

Exponential Distribution

X follows exponential distribution with parameter  $\lambda > 0$  if  $p.f.$  is given by:

f\_x(x) = { lambda \* e^(-lambda \* x) if x >= 0, 0 if x < 0 }

• Notation:  $X \sim Exp(\lambda)$

•  $E(X) = \frac{1}{\lambda}$  and  $V(X) = \frac{1}{\lambda^2}$

•  $c.d.f.$  is given by:

F\_X(x) = { 1 - e^(-lambda \* x) if x > 0, 0 if x <= 0 }

• Suppose  $X$  has exponential distribution with parameter  $\lambda > 0$ . Then for any positive numbers  $s$  and  $t$ , we have:

P(X > s + t | X > s) = P(X > t)

Normal Distribution

X follows normal distribution with mean  $\mu$  and variance  $\sigma^2$  if  $p.f.$  is given by:

f\_X(x) = 1/(sqrt(2\*pi)\*sigma) \* e^(-(x-mu)^2/(2\*sigma^2))

• Notation:  $X \sim N(\mu, \sigma^2)$

•  $E(X) = \mu$  and  $V(X) = \sigma^2$

•  $p.f.$  is bell-shaped curve and symmetric about  $x = \mu$

• Total area under curve is 1

• 2 normal curves are identical in shape if they have same  $\sigma^2$ . They differ in location by  $\mu_1 - \mu_2$ .

• As  $\sigma$  increases, curve becomes more spread out

• If  $X \sim N(\mu, \sigma^2)$  and let  $Z = \frac{X-\mu}{\sigma}$

Standardized Normal Distribution

If  $X \sim N(\mu, \sigma^2)$  and let

Z = (X - mu) / sigma

then  $Z \sim N(0, 1)$

•  $E(Z) = 0$  and  $V(Z) = 1$

•  $p.f.$  is given by:

f\_Z(z) = 1/sqrt(2\*pi) \* e^(-z^2/2)

• Importance of standardizing normal distribution is that it allows us to use tables to find probabilities

• Let  $X \sim N(\mu, \sigma^2)$ . We can compute  $P(x_1 < X < x_2)$  by standadization:

x1 < X < x2 <= (x1 - mu) / sigma < (X - mu) / sigma < (x2 - mu) / sigma

•  $c.d.f.$  is given by:

phi(z) = F\_Z(z) = integral from -infinity to z of f\_Z(z) = 1/sqrt(2\*pi) \* integral from -infinity to z of e^(-t^2/2) dt

•  $P(Z \geq 0) = P(Z \leq 0) = \phi(0) = 0.5$

• For any  $z$ ,  $\phi(z) = P(Z \leq z) = P(Z \geq -z) = 1 - \phi(-z)$

•  $-Z \sim N(0, 1)$

• If  $Z \sim N(0, 1)$ , then  $\sigma Z + \mu \sim N(\mu, \sigma^2)$

• **Quantile** - The  $\alpha$ th quantile is  $x_\alpha$  that satisfies:

P(X >= x\_alpha) = alpha

e.g. The 0.05th quantile of  $Z \sim N(0, 1)$  is 1.645, i.e.  $z_{0.05} = 1.645$ .

•  $P(Z \geq z_\alpha) = P(Z \leq -z_\alpha) = \alpha$

Normal Approximation to Binomial Distribution

Let  $X \sim B(n, p)$ , then as  $n \rightarrow \infty$ :

Z = (X - E(X)) / sqrt(V(X)) = (X - np) / sqrt(np(1-p)) ~ N(0, 1)

• Approximation is good when  $np > 5$  and  $n(1-p) > 5$

05. Sampling Distributions

Population and Sample

• Motivation: Infer something about population using sample

• **Population** - All possible observations of survey

• **Sample** - Subset of population

• Every observation can be numerical or categorical

• Finite population vs. Infinite population

Random Sampling

• Motivation: We usually know what distribution the population belongs to, but we don't know the parameters of the distribution. We can use sample to estimate the parameters.

• **Simple Random Sample** - Given sample of size  $n$ , each sample has equal chance of being selected

SRS for Infinite Population

Let  $X$  be RV with  $p.f. f_X(x)$ . Let  $X_1, X_2, \dots, X_n$  be independent random variables with same distribution as  $X$ . Then  $X_1, X_2, \dots, X_n$  is a simple random sample of size  $n$ . Joint probability function of  $X_1, \dots, X_n$ :

f\_{X1,...,Xn}(x1,...,xn) = f\_X(x1) \* f\_X(x2) \* ... \* f\_X(xn)

Sampling with Replacement

• Sampling with replacement from finite population is considered as sampling from infinite population

• Sample is random if:

• Every element in population has same probability

• Successive draws are independent

Sample Distribution of Sample Mean

• **Statistic** - Suppose random sample of  $n$  observations is  $X_1, X_2, \dots, X_n$ . A statistic is a function of  $X_1, \dots, X_n$

• **Sample Mean** -

X\_bar = 1/n \* sum from i=1 to n of X\_i

• **Sample Variance** -

S^2 = 1/(n-1) \* sum from i=1 to n of (X\_i - X\_bar)^2

• **Statistics are random variables**. We can look at distribution of statistics.

• **Sample Distribution** - Distribution of a statistic

• Mean and variance of  $\bar{X}$  :

E(X\_bar) = mu and V(X\_bar) = (sigma\_X^2) / n

Intuition:  $\mu_X$  is some unknown constant.  $\bar{X}$  guesses that. As  $n$  increases, accuracy of  $\bar{X}$  increases.

• **Standard Error** - Standard deviation of sampling distribution (e.g.  $\sigma_{\bar{X}}$ )

• **Law of Large Numbers** - As  $n$  increases,  $\bar{X}$  converges to  $\mu_X$ . i.e. For any  $\epsilon \in \mathbb{R}$ :

P(|X\_bar - mu| > epsilon) -> 0 as n -> infinity

Central Limit Theorem

If  $\bar{X}$  is mean of random sample of size  $n$  from population with mean  $\mu$  and variance  $\sigma^2$ , then as  $n \rightarrow \infty$ :

X\_bar ~ N(mu, (sigma^2)/n) approximately

• Intuition: For a large  $n$ ,  $\bar{X}$  is approximately normally distributed.

• If random sample is from normal population,  $\bar{X}$  is normally distributed no matter value of  $n$

• If very skewed, CLT may not hold even with large  $n$

Other Sampling Distributions

$\chi^2$  Distribution

• Let  $Z_1, \dots, Z_n$  be  $n$  independent and identically distributed standard normal RVs. A  $\chi^2$  RV with  $n$  degrees of freedom is defined as a RV with same distribution as  $Z_1^2 + \dots + Z_n^2$

• Notation:  $\chi^2(n)$  with  $n$  degrees of freedom

• If  $Y \sim \chi^2(n)$ , then  $E(Y) = n$  and  $V(Y) = 2n$

• For large  $n$ ,  $\chi^2(n)$  is approximately  $N(n, 2n)$

• If  $Y_1$  and  $Y_2$  are independent  $\chi^2$  RVs with  $m$  and  $n$  degrees of freedom respectively, then  $Y_1 + Y_2$  is  $\chi^2(m + n)$

•  $\chi^2$  is family of curves. All density functions have long right tail.

Sampling Distribution of  $S^2$

•  $E(S^2) = \sigma^2$

Sampling Distribution of  $\frac{(n-1)S^2}{\sigma^2}$

If  $S^2$  is variance of random sample of size  $n$  from normal population of variance  $\sigma^2$ , then:

((n-1)\*S^2) / sigma^2 = sum from i=1 to n of (X\_i - X\_bar)^2 / sigma^2

has  $\chi^2(n-1)$  distribution

t-Distribution

Suppose  $Z \sim N(0, 1)$  and  $U \sim \chi^2(n)$ . If  $Z$  and  $U$  are independent, then:

$$T = \frac{Z}{\sqrt{U/n}} \sim t(n)$$

where  $t(n)$  is called t-distribution with  $n$  degrees of freedom

- t-Distribution approaches  $N(0, 1)$  as  $n \rightarrow \infty$
- When  $n \geq 30$ , t-distribution is approximately normal
- If  $T \sim t(n)$ , then  $E(T) = 0$  and  $V(T) = \frac{n}{n-2}$  for  $n > 2$
- Symmetric about vertical axis and resembles standard normal distribution
- If  $X_1, \dots, X_n$  are independent and identically distributed normal RVs with mean  $\mu$  and variance  $\sigma^2$ , then:

$$\frac{X - \mu}{S/\sqrt{n}}$$

follows t-distribution with  $n - 1$  degrees of freedom

F-Distribution

Suppose  $U \sim \chi^2(m)$  and  $V \sim \chi^2(n)$ . If  $U_1$  and  $U_2$  are independent, then:

$$F = \frac{U/m}{V/m} \sim F(m, n)$$

is called F-distribution with  $(m, n)$  degrees of freedom

- If  $X \sim F(m, n)$ , then

$$E(X) = \frac{n}{n-2} \text{ for } n > 2$$

and

$$V(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)} \text{ for } n > 4$$

- If  $F \sim F(m, n)$ , then  $1/F \sim F(n, m)$

06. Estimation

07. Hypothesis Testing