Set Operations

1. Mutually exclusive/disjoint : $A \cap B = \varnothing$
2. $A \subset B \Rightarrow A \subseteq B$
3. $A \cap A' = \varnothing$
4. $A \cup A' = S$
5. $A \cap \varnothing = \varnothing$
6. $(A')' = A$
7. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
8. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
9. $A \cup B = A \cup (B \cap A')$
10. $A = (A \cap B) \cup (A \cap B')$
11. $(A_1 \cup A_2 \cup ... \cup A_n)' = A_1' \cap A_2' ... \cap A_n'$
12. $(A_1 \cap A_2 \cap ... \cap A_n)' = A_1' \cup A_2' ... \cup A_n'$

Counting Methods

1. $P_r^n = \frac{n!}{(n-r)!} = n(n-1)(n-2)...(n-(r-1))$
2. $\binom{n}{r} = \binom{n}{n-r} = \frac{n!}{r!(n-r)!}$

Probability

1. $0 \leq P(A) \leq 1$
2. $P(S) = 1$
3. $A \cap B = \varnothing \Rightarrow P(A \cup B) = P(A) + P(B)$
4. $P(\varnothing) = 0$
5. $P(A') = 1 - P(A)$
6. $P(A) = P(A \cap B) + P(A \cap B')$
7. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
8. $P(A \cap B) = 0 \nRightarrow A \cap B = 0$
9. $A \subset B \Rightarrow P(A) \leq P(B)$

Conditional Probability

1. $P(B|A) = \frac{P(A \cap B)}{P(A)}$
2. $P(A \cap B) = P(A)P(B|A)$
3. $P(A|B) = \frac{P(A)P(B|A)}{P(B)}$

Independence

1. $A \perp B \Leftrightarrow P(A)P(B) \Leftrightarrow P(B|A) = P(B)$

Law of Total Probability

1. $P(B) = \sum_{i=1}^{n} P(B \cap A_i) = \sum_{i=1}^{n} P(A_i)P(B|A_i)$
2. $P(B) = P(A)P(B|A) + P(A')P(B|A')$

Bayes' Theorem

1. $P(A_k|B) = \frac{P(A_k)P(B|A_k)}{\sum_{i=1}^{n} P(A_i)P(B|A_k)}$
2. $P(A|B) = \frac{P(A)P(B|A)}{P(A)P(B|A)+P(A')P(B|A')}$

PMF

1. $f(x_i) \geq 0, \forall x_i \in R_X$
2. $f(x) = 0, \forall x \notin R_X$
3. $sum_{i=1}^{\infty} f(x_i) = 1$

PDF

1. $f(x) \geq 0, \forall x_i \in R_X$ and $f(x) = 0, \forall x \notin R_X$
2. $\int_{R_x} f(x)dx = 1$
3. $P(a \leq X \leq b) = \int_a^b f(x)dx$

Discrete CDF

1. $F(x) = \sum f(t) = \sum P(X = t)$
2. $P(a \leq X \leq b) = P(X \leq b) - P(X < a) = F(b) - F(a-)$
3. $a-$:= largest value in R_X that is smaller than a

Continuous CDF

1. $F(x) = \int_{-\infty}^{x} f(t)dt$
2. $f(x) = \frac{dF(x)}{dx}$
3. $P(a \leq X \leq b) = P(a < X < b) = F(b) - F(a)$

Remarks on CDF

1. $F(x)$ is non-decreasing
2. CDF and PDF/PMF have one-to-one correspondence
3. $0 \leq F(x) \leq 1$
4. for discrete, $0 \leq f(x) \leq 1$
5. for continuous, $f(x) \geq 0$, but not necessary that $f(x) \leq 1$

Expectation

1. Discrete: $E(X) = \sum x_i f(x_i)$
2. Continuous: $E(X) = \int_{R_X} xf(x)dx$
3. $E(aX + b) = aE(X) + b$
4. $E(X + Y) = E(X) + E(Y)$
5. Discrete: $E[g(x)] = \sum g(x)f(x)$
6. Continuous: $E[g(x)] = \int_{R_X} g(x)f(x)dx$

Variance

1. $\sigma_X^2 = V(X) = E(X - \mu_X)^2$
2. Discrete: $V(X) = \sum (X - \mu_X)^2 f(x)$
3. Continuous: $V(X) = \int_{-\infty}^{\infty} (X - \mu_X)^2 f(x)dx$
4. $V(X) \geq 0$
5. $V(aX + b) = a^2 V(X)$
6. $V(X) = E(X^2) - [E(X)]^2$

Discrete Joint Probability Function

1. $f_{X,Y}(x, y) = P(X = x, Y = y)$
2. $f_{X,Y}(x, y) \geq 0, \forall (x, y) \in R_{X,Y}$
3. $f_{X,Y}(x, y) = 0, \forall (x, y) \notin R_{X,Y}$
4. $\sum_{i=1}^{\infty} \sum_{j=1}^{\infty} f_{X,Y}(x_i, y_j) = \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} P(X = x_i, Y = y_j) = 1$

Continuous Joint Probability Function

1. $P(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x, y)dydx$
2. $f_{X,Y}(x, y) \geq 0, \forall (x, y) \in R_{X,Y}$
3. $f_{X,Y}(x, y) = 0, \forall (x, y) \notin R_{X,Y}$
4. $\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} f_{X,Y}(x, y)dxdy = 1$

Marginal Probability Distribution

1. Discrete: for any x, $f_X(x) = \sum_y f_{X,Y}(x, y)$

2. Continuous: for any x, $f_X(x) = \int_{-\infty}^{\infty} f_{X,Y}(x,y)dy$
3. The intuition is to treat X as fixed variable for the **marginal probability distribution of X**

Conditional Distribution

1. **Conditional probability function of Y given** $X = x := f_{Y|X}(y|x) = \frac{f_{X,Y}(x,y)}{f_X(x)}$
2. This is the probability function of Y given x, hence it fulfills all the properties of a probability function for Y
3. $f_{X,Y}(x,y) = f_X(x)f_{Y|X}(y|x)$
4. $P(Y \le y|X = x) = \int_{-\infty}^{y} f_{Y|X}(y|x)dy$
5. $E(Y|X = x) = \int_{-\infty}^{\infty} y f_{Y|X}(y|x)dy$

Independent Random Variable

1. $f_{X,Y}(x,y) = f_X(x)f_Y(y)$
2. for any $x \in R_X$ and any $y \in R_Y$, $f_{X,Y}(x,y) = f_X(x)f_Y(y) > 0$
3. $P(X \le x; Y \le y) = P(X \le x)P(Y \le y)$
4. for arbitrary functions $g_1(\cdot)$ and $g_2(\cdot)$, $g_1(X) \perp g_2(Y)$
5. $f_{X|Y}(x|y) = f_X(x)$
6. $X \perp Y$ iff $R_x \perp R_Y$ and $f_{X,Y}(x,y) = C \times g_1(x) \times g_2(y)$
7. $f_X(x) = \frac{g_1(x)}{\sum_{t \in R_X} g_1(t)}$
8. $f_X(x) = \frac{g_1(x)}{\int_{t \in R_X} g_1(t)dt}$

Expectation and Covariance

1. $E(g(X,Y)) = \sum_x \sum_y g(x,y)f_{X,Y}(x,y)$
2. $E(g(X,Y)) = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} g(x,y)f_{X,Y}(x,y)dy\ dx$
3. $cov(X,Y) = E(XY) - E(X)E(Y)$
4. $X \perp Y \Rightarrow cov(X,Y) = 0$
5. $cov(X,Y) \not\Rightarrow X \perp Y$
6. $cov(aX + b, cY + d) = ac \cdot cov(X,Y)$
7. $V(aX + bY) = a^2V(X) + b^2V(Y) + 2ab \cdot cov(X,Y)$

Discrete Uniform Distribution

$$f_X(x) = \begin{cases} \frac{1}{k} & x = x_1, \dots x_k \\ 0 & \text{otherwise} \end{cases}$$

$$E(X) = \frac{1}{k}\sum_{i=1}^{k} x_i, \quad \sigma_X^2 = \frac{1}{k}\sum_{i=1}^{k} x_i - \mu_X^2$$

Bernoulli Random Variable

Let X be the number of success in a Bernoulli trial. Then X has only two possible values: 1 or 0

$$X \sim \text{Bernoulli}(p)$$

$$f_X(x) = p^x(1-p)^{1-x}, \text{ for } x = 0,\ 1$$

$$\mu_X = p; \quad \sigma_X^2 = pq$$

Binomial Distribution

Counts the number of successes in n trials of a Bernoulli Process.

$$X \sim \text{Binom}(n,p)$$

$$P(X = x) = \binom{n}{x} p^x(1-p)^{n-x}, \text{ for } x = 0,\ 1 \dots,\ n$$

$$E(X) = np; \quad V(X) = npq$$

Negative Binomial Distribution

Numbers of Bernoulli trials needed until the kth success occurs

$$X \sim \text{NB}(k,p)$$

$$P(X = x) = \binom{x-1}{k-1} p^k(1-p)^{x-k}, \text{ for } x = k,\ k+1,\ \dots$$

$$E(X) = \frac{k}{p}; \quad V(X) = \frac{(1-p)k}{p^2}$$

Geometric Distribution

Number of iid Bernoulli tirals needed until the first success occurs.

$$X \sim \text{Geom}(p)$$

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

$$E(X) = \frac{1}{p}; \quad V(X) = \frac{(1-p)}{p^2}$$

Poisson Distribution

Number of events occurring in a fixed period of time or fixed region.

$$X \sim \text{Poisson}(\lambda)$$

$$P(X = x) = \frac{e^{-\lambda}\lambda^k}{k!}$$

$$E(X) = \lambda; \quad V(X) = \lambda$$

Poisson Approximation to Binomial

Let $X \sim \text{Binom}(n,p)$. If $n \to \infty$ and $p \to 0$ s.t. $\lambda = np$, then $X \sim \text{Poisson}(np)$

$$\lim_{p \to 0; n \to \infty} P(X = x) = \frac{e^{-np}(np)^x}{x!}$$

$$n \ge 20 \text{ and } p \le 0.05 \text{ or } n \ge 100 \text{ and } np \le 10$$

Continuous Uniform Distribution

$$X \sim U(a,b)$$

$$f_X(x) = \begin{cases} \frac{1}{b-1} & a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

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

$$F_X(x) = \begin{cases} 0 & x < a \\ \frac{x-a}{b-a} & a \le x \le b \\ 1 & x > b \end{cases}$$

Exponential Distribution

Often used to model waiting time to the first success in continuous time.

$$X \sim \text{Exp}(\lambda)$$

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

$$E(X) = \frac{1}{\lambda}, \quad V(X) = \frac{1}{\lambda^2}$$

$$F_X(x) = \begin{cases} 1 - e^{-\lambda x} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

$$P(X > x) = e^{-\lambda x}, \text{ for } x > 0$$

$$f_X(x) = \begin{cases} \frac{1}{\mu}e^{-x/\mu} & x \ge 0 \\ 0 & x < 0 \end{cases}$$

$$E(X) = \mu, \quad V(X) = \mu^2, \quad F_X(x) = 1 - e^{-x/\mu}$$

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

Normal Approximation to Binomial

If $n \to \infty$ and p remains a constant.

$$Z = \frac{X - np}{\sqrt{np(1-p)}} \sim N(0,1)$$

$$np < 5 \text{ and } n(1-p) > 5$$

Continuity Correction

Apply continuity correction when approximating the binomial using the normal.

1. $P(X = k) \approx P(k - 1/2 < X < k + 1/2)$
2. $P(a \leq X \leq b) \approx P(a - 1/2 < X < b + 1/2)$
3. $P(a < X \leq b) \approx P(a + 1/2 < X < b + 1/2)$
4. $P(a \leq X < b) \approx P(a - 1/2 < X < b - 1/2)$
5. $P(a < X < b) \approx P(a + 1/2 < X < b - 1/2)$
6. $P(X \leq c) = P(0 \leq X \leq c) \approx P(-1/2 < X < c + 1/2)$
7. $P(X > c) = P(c < X \leq n) \approx P(c{+}1/2 < X < n{+}1/2)$

Sample mean: $\bar{x} = \frac{1}{n}\sum_{i=1}^n x_i$

Sample variance: $x^2 = \frac{1}{n-1}\sum_{i=1}^n (x_i - \bar{x})^2$

$$\mu_{\bar{X}} = E(\bar{X}), \quad \sigma^2_{\bar{X}} = V(\bar{X}) = \frac{\sigma^2_X}{n}$$

χ^2 Distribution

Let $Z_1, \dots, Z_n$ be n iid standard normal RV, An RV with the same distribution as $Z_1^2 + \cdots + Z_n^2$ is a χ^2 RV with n dof.

1. $Y \sim \chi^2(n)$
2. $E(Y) = n$ and $V(Y) = 2n$
3. For large n, $\chi^2(n)$ is approximately $N(n, 2n)$
4. If Y_1 and Y_2 are χ^2 RV with m and n dof respectively, then $Y_1 + Y_2$ is a χ^2 RV with $m + n$ dof
5. $\frac{(n-1)S^2}{\sigma^2} = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{\sigma^2}$ has a χ^2 dist with $n-1$ dof

t-Distribution

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

$$T = \frac{Z}{\sqrt{U/n}}$$

follows the t-distribution with n dof

1. If $n \to \infty$, t-dist approaches $N(0,1)$. When $n \geq 30$, we can replace it by $N(0,1)$.
2. If $T \sim t(n)$, then $E(T) = 0$ and $V(T) = n/(n-2)$ for $n > 2$

If $X_1, \dots, X_n$ are iid normal RV with mean μ and variance σ^2, then

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

follows a t-dist. with $n-1$ dof

F-Distribution

Suppose $U \sim \chi^2(m) \perp V \sim \chi^2(n)$, then the distribution of the RV

$$F = \frac{U/m}{V/n}$$

is a F-distribution with (m, n) dof

1. $E(X) = \frac{n}{n-2}$
2. $V(X) = \frac{2n^2(m+n-2)}{m(n-2)^2(n-4)}$
3. If $F \sim F(n, m)$ then $1/F \sim F(m, n)$
4. $F(m, n; 1-\alpha) = 1/F(n, m; \alpha)$

Suppose that random samples of sizes n_1 and n_2 are selected from two normal populations with variances σ_1^2 and σ_1^2 respectively.

$$U = \frac{(n_1 - 1)S_1^2}{\sigma_1^2} \sim \chi^2(n_1 - 1)$$

$$V = \frac{(n_2 - 1)S_2^2}{\sigma_2^2} \sim \chi^2(n_2 - 1)$$

Therefore we have

$$F = \frac{U/(n_1 - 1)}{V/(n_2 - 1)} = \frac{S_1^2/\sigma_1^2}{S_2^2/\sigma_2^2} \sim F(n_1 - 1, n_2 - 1)$$

Maximum Error of Estimate

$$E = z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$$

$$P(\frac{|\bar{X} - \mu|}{\sigma/\sqrt{n}} \leq z_{\alpha/2}) = P(|\bar{X} - \mu| \leq z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}) = 1 - \alpha$$

$$n \geq (\frac{z_{\alpha/2} \cdot \sigma}{E_0})^2$$

Population	σ	n	Statistic	E
Normal	known	any	$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	$z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$
any	known	large	$Z = \frac{\bar{X}-\mu}{\sigma/\sqrt{n}}$	$z_{\alpha/2} \cdot \frac{\sigma}{\sqrt{n}}$
Normal	unknown	small	$T = \frac{\bar{X}-\mu}{S/\sqrt{n}}$	$t_{n-1;\alpha/2} \cdot \frac{S}{\sqrt{n}}$
any	unknown	large	$Z = \frac{\bar{X}-\mu}{S/\sqrt{n}}$	$z_{\alpha/2} \cdot \frac{S}{\sqrt{n}}$

Independent Samples (Known and Unequal Variances)

If two populations are normal OR both samples are large,

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}}$$

$$(\bar{X} - \bar{Y}) \pm z_{\alpha/2}\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}$$

Independent Samples (Large, Unknown and Unequal Variances)

$$S^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \bar{X})^2$$

$$Z = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

$$(\bar{X} - \bar{Y}) \pm z_{\alpha/2}\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}$$

Independent Samples (Small, Unknown but Equal Variances)

Assume equal variances if $1/2 \leq S_1/S_2 \leq 2$

$$S_p^2 = \frac{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2}{n_1 + n_2 - 2}$$

$$T = \frac{(\bar{X} - \bar{Y}) - (\mu_1 - \mu_2)}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

$$(\bar{X} - \bar{Y}) \pm t_{n_1+n_2-2;\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Independent Samples (Large, Unknown but Equal Variances)

$$(\bar{X} - \bar{Y}) \pm z_{\alpha/2} S_p \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}$$

Paired Data

- X_i and Y_i are dependent
- (X_i, Y_i) and (X_j, Y_j) are independent for any $i \neq j$
- Define $D_i = X_i - Y_i$, $\mu_D = \mu_1 - \mu_2$

$$T = \frac{\bar{D} - \mu_D}{S_D/\sqrt{n}}$$

$$\bar{D} = \frac{\sum_{i=1}^{n} D_i}{n}, \quad S_D^2 = \frac{\sum_{i=1}^{n}(D_i - \bar{D})^2}{n-1}$$

If $n < 30$, $T \sim t_{n-1}$, CI = $\bar{d} \pm t_{n-1;\alpha/2} \cdot \frac{S_D}{\sqrt{n}}$
If $n \geq 30$, $T \sim N(0,1)$, CI = $\bar{d} \pm z_{\alpha/2} \cdot \frac{S_D}{\sqrt{n}}$

Type I & Type II Error

- Type I Error: ejecting H_0 when $H_0$is True
- type II Error: Not rejecting H_0 when H_0 is False
- $\alpha = P(Type\ I\ Error) = P(Reject\ H_0|H_0\ is\ true)$
- $\beta = P(Type\ II\ Error) = P(Do\ Not\ Reject\ H_0|H_0\ is\ false)$
- Power of the test: 1-β = P(Reject H_0—H_0 is false)

Hypothesis Test: Known Variance

- population variance σ^2 is known
- underlying distribution is normal
- n is large

Test statistics:

$$Z = \frac{\bar{X} - \mu_0}{\sigma/\sqrt{n}} \sim N(0,1)$$

Hypothesis Test: Unknown Variance

- population variance σ^2 is unknown
- underlying distribution is normal

Test statistics:

$$T = \frac{\bar{X} - \mu_0}{S/\sqrt{n}} \sim t_{n-1}$$

when $n \geq 30$, we can replace t_{n-1} by Z

Independent Samples

1. - σ_1^2 and σ_2^2 are known
   - n_1 and n_2 are large

$$Z = \frac{(\bar{X} - \bar{Y}) - \delta_0}{\sqrt{\frac{\sigma_1^2}{n_1} + \frac{\sigma_2^2}{n_2}}} \sim N(0,1)$$

2. - σ_1^2 and σ_2^2 are unknown
   - n_1 and n_2 are large

$$Z = \frac{(\bar{X} - \bar{Y}) - \delta_0}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}} \sim N(0,1)$$

3. - σ_1^2 and σ_2^2 are unknown but equal
   - underlying distribution is normal
   - n_1 and n_2 are small

$$Z = \frac{(\bar{X} - \bar{Y}) - \delta_0}{S_p\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \sim t_{n_1+n_2-2}$$

Paired Data

$$T = \frac{\bar{D} - \mu_{D_0}}{S_D/\sqrt{n}}$$

If n ¡ 30 and the population is normally distributed, then

$$T \sim t_{n-1}$$

If n $\geq$ 30, then

$$T \sim N(0,1)$$

Inequalities and Absolute Values

1. $|x| < a \iff -a < x < a$
2. $|x| > a \iff x < -a$ or $x > a$