Probability Theory and Mathematical Statistics

I. Probability Formulas

- **Addition Formula**: $P(A+B) = P(A \cup B) = P(A) + P(B) - P(AB)$
- **Subtraction Formula**: $P(A-B) = P(A\overline{B}) = P(A) - P(AB)$
- **Conditional Probability**: $P(A|B) = \dfrac{P(AB)}{P(B)}$, $P(B) > 0$
- **Compatible Events**: $P(AB) > 0$
- **Mutually Exclusive Events**: $P(AB) = 0$
- **Independent Events**: $P(AB) = P(A)P(B)$
- **Distribution Function**:
  - $F(a) = P\{X \le a\}$
  - $P\{X < a\} = \lim\limits_{x \to a^-} F(x)$
  - $f_X(x) = \dfrac{\mathrm{d}F_X(x)}{\mathrm{d}x}$
- **Convolution Function** $(Z = X + Y)$: $f_Z(z) = \int_{-\infty}^{\infty} f_X(x) f_Y(z-x)\,\mathrm{d}x$
- **Convolution Function** $(Z = X - Y)$: $f_Z(z) = \int_{-\infty}^{\infty} f_X(x)\, f_Y(x-z)\,\mathrm{d}x$

II. Numerical Characteristics

- **Mathematical Expectation**:
  - Continuous Type: If the probability density of X is $f(x)$, then $E(X) = \displaystyle\int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x$
  - Function of Random Variable: $E[g(X)] = \displaystyle\int_{-\infty}^{+\infty} g(x) f(x)\,\mathrm{d}x$
- **Marginal Expectation in Two-Dimensional Case**: If the joint probability density of (X, Y) is $f(x, y)$, its marginal probability densities are:

$$f_X(x) = \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}y$$
$$f_Y(y) = \int_{-\infty}^{+\infty} f(x, y)\,\mathrm{d}x$$

- **Variance**:
  - Definition: $D(X) = E[(X - E(X))^2]$

- Computational Formula: $D(X) = E(X^2) - [E(X)]^2$
- Property: $D(CX) = C^2 D(X)$

- **Correlation Coefficient**: $\rho_{XY} = \dfrac{\mathrm{Cov}(X,Y)}{\sqrt{D(X)} \cdot \sqrt{D(Y)}}$

- **Equivalent Propositions of Uncorrelated**:

$$\mathrm{Cov}(X,Y) = 0 \iff \rho_{XY} = 0 \iff E(XY) = E(X)E(Y) \iff D(X+Y) = D(X)+D(Y)$$

- **Covariance**:
  - Definition: $\mathrm{Cov}(X,Y) = E[(X - E(X))(Y - E(Y))]$
  - Computational Formula: $\mathrm{Cov}(X,Y) = E(XY) - E(X)E(Y)$

- **Properties of Covariance**
  1. $\mathrm{Cov}(X,Y) = \mathrm{Cov}(Y,X)$
  2. $\mathrm{Cov}(X,C) = 0$
  3. $\mathrm{Cov}(aX,bY) = ab\,\mathrm{Cov}(X,Y)$
  4. $\mathrm{Cov}(X,X) = D(X)$
  5. $\mathrm{Cov}(X+Y,Z) = \mathrm{Cov}(X,Z) + \mathrm{Cov}(Y,Z)$
  6. $D(X \pm Y) = D(X) + D(Y) \pm 2\,\mathrm{Cov}(X,Y)$
  7. X, Y are independent $\Rightarrow \mathrm{Cov}(X,Y) = 0$

III. Probability Distributions

- **Poisson Distribution**: $X \sim P(\lambda)$
  - Conclusion: If $X \sim P(\lambda_1)$, $Y \sim P(\lambda_2)$, and X and Y are independent, then $X + Y \sim P(\lambda_1 + \lambda_2)$

- **Normal Distribution**: $X \sim N(\mu, \sigma^2)$
  - Standardization and Probability:

$$P\{a < X \le b\} = P\left\{\frac{a-\mu}{\sigma} < \frac{X-\mu}{\sigma} \le \frac{b-\mu}{\sigma}\right\} = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$$

  - Conclusion: If $Z \sim N(0,1)$, then

$$\forall a > 0, \quad P\{|Z| \le a\} = 2\Phi(a) - 1$$

- **Exponential Distribution**: $X \sim E(\lambda)$ $(\lambda > 0)$
  - Conclusion:
    * $P\{X > a\} = e^{-\lambda a} \quad (a > 0)$
    * $P\{X > s + t \mid X > s\} = P\{X > t\}$, where $s, t > 0$

- **Common Distributions Table**:

Type	Notation	Distribution Law/Density	Expectation	Variance
0-1	$X \sim b(1,p)$	$P\{X=k\} = p^k(1-p)^{1-k} \quad k=0,1$	p	$p(1-p)$
Binomial	$X \sim B(n,p)$	$P\{X=k\} = C_n^k p^k (1-p)^{n-k} k=0,1,\ldots,n$	np	$np(1-p)$
Poisson	$X \sim \pi(\lambda)$	$P\{X=k\} = \frac{\lambda^k}{k!}e^{-\lambda} \quad k=0,1,2,\ldots$	λ	λ
Uniform	$X \sim U(a,b)$	$f(x) = \begin{cases} \frac{1}{b-a}, & a<x<b \\ 0, & \text{otherwise} \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Normal	$X \sim N(\mu,\sigma^2)$	$f(x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad -\infty<x<+\infty$	μ	σ^2
Exponential	$X \sim E(\lambda)$	$f(x) = \begin{cases} \lambda e^{-\lambda x}, & x>0 \\ 0, & x \le 0 \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$

IV. Mathematical Statistics

- **Common Statistics**:
  - Sample Mean: $\overline{X} = \dfrac{1}{n}\sum_{i=1}^{n} X_i$
  - Sample Variance: $S^2 = \dfrac{1}{n-1}\sum_{i=1}^{n}(X_i - \overline{X})^2$
- **Three Major Sampling Distributions**:
  - χ^2 **Distribution**:
    * Definition: $X_1, \ldots, X_n \overset{\text{i.i.d.}}{\sim} N(0,1)$, then $\sum_{i=1}^{n} X_i^2 \sim \chi^2(n)$
    * Property: $E(\chi^2) = n,\ D(\chi^2) = 2n$
  - t **Distribution**:
    * Definition: $X \sim N(0,1), Y \sim \chi^2(n)$ independent, then $T = \dfrac{X}{\sqrt{Y/n}} \sim t(n)$
    * Property: Probability density $f(t)$ is an even function; if $T \sim t(n)$, then $T^2 \sim F(1,n)$
  - F **Distribution**:
    * Definition: $X \sim \chi^2(n_1), Y \sim \chi^2(n_2)$ independent, then $F = \dfrac{X/n_1}{Y/n_2} \sim F(n_1,n_2)$
    * Property: If $F \sim F(n_1,n_2)$, then $\dfrac{1}{F} \sim F(n_2,n_1)$
- **Properties of Statistics**: Let $X_1, X_2, \ldots, X_n$ be a sample from population X, with $E(X) = \mu,\ D(X) = \sigma^2$, then:
  - $E(X_i) = \mu,\ D(X_i) = \sigma^2,\ E(\overline{X}) = \mu,\ D(\overline{X}) = \dfrac{\sigma^2}{n}$

- **Testing Methods for Mean and Variance of Normal Population**:

Null Hypothesis H_0	Test Statistic	Rejection Region
$\mu \leq \mu_0$	$Z = \dfrac{\overline{X} - \mu_0}{\sigma/\sqrt{n}}$	$z \geq z_\alpha$
$\mu = \mu_0$		$\lvert z\rvert \geq z_{\alpha/2}$
$\mu \geq \mu_0$		$z \leq -z_\alpha$
$\mu \leq \mu_0$	$t = \dfrac{\overline{X} - \mu_0}{S/\sqrt{n}}$	$t \geq t_\alpha(n-1)$
$\mu = \mu_0$		$\lvert t\rvert \geq t_{\alpha/2}(n-1)$
$\mu \geq \mu_0$		$t \leq -t_\alpha(n-1)$
$\sigma^2 \leq \sigma_0^2$	$\chi^2 = \dfrac{(n-1)S^2}{\sigma_0^2}$	$\chi^2 \geq \chi^2_\alpha(n-1)$
$\sigma^2 = \sigma_0^2$		$\chi^2 \leq \chi^2_{1-\alpha/2}(n-1)$ or $\chi^2 \geq \chi^2_{\alpha/2}(n-1)$
$\sigma^2 \geq \sigma_0^2$		$\chi^2 \leq \chi^2_{1-\alpha}(n-1)$

V. Parameter Estimation and Hypothesis Testing

- **Unbiasedness**: If $E(\hat{\theta}) = \theta$, then $\hat{\theta}$ is called an unbiased estimator of θ.
- **Efficiency**: If $\hat{\theta}_1$ and $\hat{\theta}_2$ are both unbiased estimators of θ, and $D(\hat{\theta}_1) < D(\hat{\theta}_2)$, then $\hat{\theta}_1$ is more efficient than $\hat{\theta}_2$.
- **Method of Moments**: $E(X) = \overline{X}$
  - $\int_{-\infty}^{+\infty} x f(x)\,\mathrm{d}x = \frac{1}{n}\sum_{i=1}^{n} X_i$
- **Steps of Maximum Likelihood Estimation**:
  - $L(\theta) = \displaystyle\prod_{i=1}^{n} p(x_i;\theta) \Rightarrow \ln L(\theta) = \sum_{i=1}^{n} \ln p(x_i;\theta) \Rightarrow \frac{\mathrm{d}[\ln L(\theta)]}{\mathrm{d}\theta} = 0$
- **Confidence Interval for μ of Single Normal Population** (Confidence Level $1-\alpha$):
  1. σ^2 known: $\left(\overline{X} - Z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}},\ \overline{X} + Z_{\alpha/2}\cdot\dfrac{\sigma}{\sqrt{n}}\right)$
  2. σ^2 unknown: $\left(\overline{X} - t_{\alpha/2}(n-1)\cdot\dfrac{S}{\sqrt{n}},\ \overline{X} + t_{\alpha/2}(n-1)\cdot\dfrac{S}{\sqrt{n}}\right)$
- **Two Types of Errors in Hypothesis Testing**:
  - **Type I Error (Rejecting Truth)**: When H_0 is true, H_0 is rejected.
  - **Type II Error (Accepting False)**: When H_0 is false, H_0 is accepted.