

## C1: Probability

### 3 Kolmogorov's Axioms of Probability

1.  $\forall$  event  $E(P(E) \in [0, 1])$  | 2.  $P(\xi) = 1$  | 3.  $\forall i \neq j (E_i \cap E_j = \emptyset) \Rightarrow P(\bigcup_{\alpha=1}^{\infty} E_{\alpha}) = \sum_{\alpha=1}^{\infty} P(E_{\alpha})$ .

Else generally for Axiom 3, Boole's Inequality ( $\because$  PIE) : LHS  $\leq$  RHS

### Frequentist vs Bayesian Interpretation

Frequentist (ST2132):  $P(E)$  = limiting relative freq in  $\infty$  trials aggregated as mass phenomena E

Bayesian: posterior  $\pi(\theta|E) = \frac{P(E|\theta)\pi(\theta)}{P(E)}$  for new evidence E updating prior  $= \pi(\theta)$ ,  $P(E) = \int P(E|\theta)\pi(\theta)d\theta$

## C2: Expectation

### Moments

$$\text{Raw (crude) moment (abt origin)} \mu_k = E[X^k] = \begin{cases} \int_{-\infty}^{\infty} x^k f_X(x) dx & \text{if cont X} \\ \sum_i x_i^k P(X = x_i) & \text{if disc X} \end{cases}$$

$$\text{Central moment (abt mean)} \mu'_k = E[(X - E[X])^k] = \begin{cases} \int_{-\infty}^{\infty} (x - \mu)^k f_X(x) dx & \text{if cont X} \\ \sum_i (x - \mu)^k P(X = x_i) & \text{if disc X} \end{cases} \mid \text{Standardised moment } \hat{\mu}_k = \frac{\mu'_k}{\sigma^k}$$

$\therefore$  population mean  $\mu = \mu_1$ , population variance  $\text{Var}(X) := E[(X - \mu)^2] = \mu_2 - \mu_1^2 \geq 0$ , population skewness  $\gamma = \hat{\mu}_3$

### Expectation Properties

Linearity of Expectation:  $E[aX + bY + c] = aE[X] + bE[Y] + c|E[X] = \int_0^{\infty} (P(X > x)dx - P(X < -x))dx (\because \text{Fubini Thm})$

Markov Inequality: If  $X > 0$  cont,  $\forall a > 0$ ,  $E[X] \geq aP(X \geq a)$  | Chebyshev Inequality:  $P(|X - \mu| \geq a) \leq \frac{E[(X - \mu)^2]}{a^2} = \frac{\sigma^2}{a^2}$

1D Change-of-Variables Formula: If  $Y = g(X)$  cont monotonous,  $f_Y(y) = f_X(g^{-1}(y)) \left| \frac{dy}{dx} g'(y) \right| (\because f_Y(y) = f_X(g^{-1}(y)))$

$\therefore$  Generally, Jacobian:  $f_Y(y)dy = f_X(g^{-1}(y))|J_{g^{-1}}(y)| \Leftrightarrow \prod dy_i = |\frac{\partial(y_1, \dots, y_n)}{\partial(x_1, \dots, x_n)}| \prod dx_i = |\begin{pmatrix} y_{1x_1} & \cdots & y_{1x_n} \\ \vdots & \ddots & \vdots \\ y_{nx_1} & \cdots & y_{nx_n} \end{pmatrix}| \prod dx_i$

Law of the Unconscious Statistician (LOTUS):  $E[g(X)] = \int_{-\infty}^{\infty} g(x)f_X(x)dx (\therefore \text{independent of } f_g(x)(g(x)))$

## C3: Variance & Multinomial Distribution

### Covariance & Variance

$\text{Cov}(X, Y) = E[(X - E[X])(Y - E[Y])] = E[XY] - E[X]E[Y]$  |  $\text{Cov}(X, X) = \text{Var}(X)|\text{Cov}(-X, Y) = \text{Cov}(X, -Y) = -\text{Cov}(X, Y)$

$\text{Cov}(W, aX + bY + c) = a\text{Cov}(W, X) + b\text{Cov}(W, Y) \Rightarrow \text{Var}(aX + bY + c) = a^2\text{Var}(X) + b^2\text{Var}(Y) + 2ab\text{Cov}(X, Y) (\therefore \text{Ac})$

Population Pearson correlation  $\rho_{X,Y} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y} \in [-1, 1] \Rightarrow \text{Sample Pearson correlation } r_{X,Y} = \frac{\sum x_i y_i - \bar{x}\bar{y}}{(n-1)s_X s_Y}$

### Non-correlation & Independence

$X, Y$  uncorrelated  $\overset{\text{def}}{\Leftrightarrow} \text{Cov}(X, Y) = \rho_{X,Y} = 0 \Leftrightarrow E[XY] = E[X]E[Y] \Leftrightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} xy f_X(x) f_Y(y) dxdy = (\int_{-\infty}^{\infty} x f_X(x) dx)(\int_{-\infty}^{\infty} y f_Y(y) dy)$

$X, Y$  independent  $\overset{\text{def}}{\Leftrightarrow} f_{XY} = f_X f_Y \Rightarrow X, Y$  uncorrelated ( $\because$  Fubini Thm). Converse non-ex:  $X \sim \text{Uniform}(-\frac{1}{2}, \frac{1}{2}), Y \sim X^2$

### Joint CDF & Marginal PDF

Joint CDF  $F(t_1, \dots, t_k) = P(\{X_i \leq t_i : i \in [1, k]\})$ . Marginal PDF  $f_{X_i}(t_i) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} f(t_1, \dots, t_k) \prod_{j \neq i} dt_j$

If  $k = 2$ ,  $P(\{\ell_1 \leq X_i \leq h_i : i \in [1, 2]\}) = F(h_1, h_2) + F(\ell_1, \ell_2) - F(\ell_1, h_2) - F(h_1, \ell_2)$  ( $\because$  PIE)

Multinomial discrete RV  $\vec{X} \sim \text{Multinomial}(N, p_1, \dots, p_k), \sum_{i=1}^k p_i = 1, \sum_{i=1}^k X_i = N$  for k disjoint events

$$f_{\vec{X}}(\vec{x}) = \binom{N}{x_1, \dots, x_k} \prod_{i=1}^k p_i^{x_i} = \frac{N!}{x_1! \dots x_k!} \prod_{i=1}^k p_i^{x_i} \therefore \sum_{\sum x_i = 1, x_i \geq 0} \binom{N}{x_1, \dots, x_k} \prod_{i=1}^k p_i^{x_i} = 1$$

For N independent trials,  $E[X_i] = Np_i \Rightarrow E[\vec{X}] = N\vec{p}$ ,  $\text{Var}(X_i) = Np_i(1-p_i)$

$$\therefore \forall i \neq j \text{ Cov}(X_i, X_j) = E[X_i X_j] - E[X_i]E[X_j] = E[\sum_{r=1}^N I_r^{(i)} \sum_{s=1}^N I_s^{(j)}] - N^2 p_i p_j (\because I_r^{(i)} = \text{indicator if rth trial} = X_i) \\ = \sum_{r=s} E[\sum_{r=1}^N I_r^{(i)} \sum_{s=1}^N I_s^{(j)}] + \sum_{r \neq s} E[\sum_{r=1}^N I_r^{(i)} \sum_{s=1}^N I_s^{(j)}] - N^2 p_i p_j \\ = 0 + N(N-1)p_i p_j - N^2 p_i p_j = -Np_i p_j < 0$$

$\therefore \text{Var}(\vec{X}) = N(\text{diag}(\vec{p}) - \vec{p}\vec{p}^T)$  p.s.d. symmetric in  $M_{k \times k}(\mathbb{R})$  | Moment-generating function  $m(\vec{X}) = E[e^{\vec{t}^T \vec{X}}] = (\sum p_i e^{t_i})^N$

Binomial discrete RV  $X \sim \text{Binomial}(N, p)$  ( $\because k = 2$  from multinomial)

$f_X(x) = P(X = x) = \binom{N}{x} p^x (1-p)^{N-x} \therefore \sum_{x=0}^N \binom{N}{x} p^x (1-p)^{N-x} = 1 | m(X) = E[e^{tX}] = (1-p + pe^t)^N$  ( $\because t_2 = 0$  omit)

$E[X] = Np, E[X^2] = N(N-1)p^2 + Np, E[X^3] = \prod_{i=0}^2 (N-i)p^3 + 3N(N-1)p^2 + Np, \text{Var}(X) = Np(1-p), \gamma(X) = \frac{1-2p}{Np(1-p)}$

$$E[X^4] = \prod_{i=0}^3 (N-i)p^4 + 6 \prod_{i=0}^2 (N-i)p^3 + 7N(N-1)p^2 + Np, \text{Mode} = \begin{cases} \lfloor (N+1)p \rfloor & \forall (N+1)p \in \mathbb{Z}^c \cup \{0\} \\ \binom{N+1}{2}p, (N+1)p-1 & \forall (N+1)p \in \{1, \dots, N\} \\ N & \text{if } (N+1)p = N+1 \end{cases}$$

## C4: Discrete & Continuous Distributions

Bernoulli discrete RV  $X \sim \text{Bernoulli}(p) \equiv \text{Binomial}(1, p)$

$$f_X(x) = \begin{cases} p & \text{if } X = 1 \\ 1-p & \text{if } X = 0 \end{cases} \quad F_X(x) = \begin{cases} 0 & \forall X \leq 0 \\ 1-p & \forall X \in [0, 1] \\ 1 & \forall X \geq 1 \end{cases} \quad \forall n \in \mathbb{N}^+ (E[X^n] = p). \quad \text{Var}(X) = p(1-p), m(X) = 1-p + pe^t$$

Poisson discrete RV  $X \sim \text{Poisson}(\lambda) \equiv \lim_{N \rightarrow \infty, p \rightarrow 0} \text{Binomial}(N, p), \lambda \rightarrow Np$

$$f_X(x) = e^{-\lambda} \frac{\lambda^x}{x!} \therefore \sum_0^{\infty} \frac{\lambda^x}{x!} = e^{\lambda}. \quad F_X(x) = e^{-\lambda} \sum_{i=0}^{\lfloor x \rfloor} \frac{\lambda^i}{i!} = \frac{\Gamma(\lfloor x+1 \rfloor, \lambda)}{\lfloor x \rfloor!}. \quad \text{Mode} = \begin{cases} \lfloor \lambda \rfloor & \forall \lambda \notin \mathbb{N}^+ \\ \lambda, \lambda-1 & \forall \lambda \in \mathbb{N}^+ (\because \frac{f_X(i)}{f_X(i-1)} = \frac{\lambda}{i-1}) \end{cases}$$

$\therefore \forall k \in \mathbb{N}^+ (E[\prod_{i=0}^k (X-i)] = \lambda^k) \Leftrightarrow E[X^n] = \sum_{i=0}^n \lambda^i S(n, i)$  with Stirling's 2nd kind

$\therefore E[X] = \lambda, E[X^2] = \lambda^2 + \lambda, E[X^3] = \lambda^3 + 3\lambda^2 + \lambda, E[X^4] = \lambda^4 + 6\lambda^3 + 7\lambda^2 + \lambda. \quad \text{Var}(X) = \lambda, \gamma(X) = \lambda^{-\frac{1}{2}}$

Negative Binomial discrete RV  $X \sim \text{NegBinom}(r, p)$ : trials (not failures) to r successes each of prob p

$$f_X(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r} \therefore \sum_{x=r}^{\infty} \binom{x-1}{r-1} p^r (1-p)^{x-r} = 1. \quad \forall i \in \mathbb{N}^+ \frac{f_X(i)}{f_X(i-1)} = \frac{p(i-1)}{(1-p)(i-r)} \Rightarrow \text{Mode} = \lfloor \frac{r-(r+1)p}{1-2p} \rfloor$$

$$E[X] = \frac{r}{p}, E[X^2] = \frac{r(r+1)}{p^2} - \frac{r}{p}, E[X^3] = \frac{r(r+1)(r+2)}{p^3} - \frac{3r(r+1)}{p^2} + \frac{4r}{p}, \text{Var}(X) = \frac{r}{p}(\frac{1}{p}-1), m(X) = [\frac{pe^t}{1+(p-1)e^t}]^r \forall t < -\ln(1-p)$$

## Geometric discrete RV X ~ Geometric(p) $\equiv \text{NegBinom}(1, p)$

$$f_X(x) = (1-p)^{x-1} p, E[X] = \frac{1}{p}, E[X^2] = \frac{2}{p^2} - \frac{1}{p}, E[X^3] = \frac{6}{p^3} - \frac{6}{p^2} + \frac{4}{p}, \text{Var}(X) = \frac{1}{p}(\frac{1}{p}-1), \text{Mode} = 1$$

## Hypergeometric discrete RV X ~ Hypergeometric(N, m, n): $\checkmark$ in n-sample in N-pop with m✓ w/o replacement

$$f_X(x) = \frac{\binom{N-m}{n-x}}{\binom{N}{n}} \therefore \sum_{x=0}^{\min\{m, n\}} \frac{\binom{N-m}{n-x}}{\binom{N}{n}} = 1 (\because \text{Vandermonde Identity}), E[X] = \frac{mn}{N}, E[X^2] = \frac{m(m-1)n(n-1)}{N(N-1)}$$

$$\text{Var}(X) = \frac{mn}{N} \frac{(N-m)(N-n)}{N(N-1)}, m(X) = \frac{(N-m)}{N} \frac{\Gamma(m-N+1)}{\Gamma(n) \Gamma(m+n-N+1)} \int_0^1 \frac{t^{-n-1} (1+t)^{m+n-N}}{(1-tz)^N} dt \leq m(X)_{\text{binom}}$$

if same mean

## Negative Hypergeometric discrete RV X ~ NegHypergeom(N, m, r) w/o replacement

$$f_X(x) = \frac{\binom{m}{r-1} \binom{N-m}{x-r}}{\binom{N}{x-1}} \therefore \sum_{x=1}^{\infty} \frac{\binom{m}{r-1} \binom{N-m}{x-r}}{\binom{N}{x-1}} = \frac{m}{N}$$

Take  $Y = X - r$ ,  $E[X] = \frac{r(N+1)}{m+1}$ ,  $\text{Var}(X) = \text{Var}(Y) = \frac{(N-m)r(m+1-r)(N+1)}{(m+1)^2(m+2)}$

## Uniform continuous RV X ~ Uniform(α, β)

$$f_X(x) = \begin{cases} \frac{1}{\beta-\alpha} & \forall x \in [\alpha, \beta] \\ 0 & \text{elsewhere} \end{cases} \quad F_X(x) = \begin{cases} 0 & \forall x \leq \alpha \\ \frac{x-\alpha}{\beta-\alpha} & \forall x \in [\alpha, \beta] \\ 1 & \forall x \geq \beta \end{cases} \quad \text{Var}(X) = \frac{(\beta-\alpha)^2}{12}, \gamma(X) = 0, m(X) = \frac{e^{t\beta}-e^{t\alpha}}{t(\beta-\alpha)}$$

$\therefore 68-95-99$  Rule:  $\text{erf}(\frac{1}{\sqrt{2}}) = 0.682689, \text{erf}(\sqrt{2}) = 0.954500, \text{erf}(\frac{3}{\sqrt{2}}) = 0.997300$

$$\text{erf}^{-1}(0.9) = 1.16309, \text{erf}^{-1}(0.95) = 1.38590, \text{erf}^{-1}(0.975) = 1.58491, \text{erf}^{-1}(0.99) = 1.82139, \text{erf}^{-1}(0.999) = 2.32675$$

$$\text{Error Function } \text{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt = \frac{2}{\sqrt{\pi}} \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{n!(2n+1)} \text{ odd. } \text{erf}(\infty) = 1 \Leftrightarrow \int_{-\infty}^{\infty} e^{-t^2} dt = \sqrt{\pi}, \text{erf}(0) = 0$$

$\therefore$  68-95-99 Rule:  $\text{erf}(\frac{1}{\sqrt{2}}) = 0.682689, \text{erf}(\sqrt{2}) = 0.954500, \text{erf}(\frac{3}{\sqrt{2}}) = 0.997300$

$$\text{Normal (Gaussian) continuous RVX} \sim \text{Normal}(\mu, \sigma^2)$$

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad F_X(x) = \int_{-\infty}^x f_X dx = \frac{1}{2} + \frac{1}{2} \text{erf}(\frac{x-\mu}{\sqrt{2\sigma^2}}) \quad E[X] = \text{Med} = \text{Mode} = \mu, E[X^2] = \mu^2 + \sigma^2$$

$$E[X^3] = \mu^3 + 3\mu\sigma^2, E[X^4] = \mu^4 + 6\mu^2\sigma^2 + 3\sigma^4, \text{Var}(X) = \sigma^2, m(X) = e^{\frac{1}{2}\sigma^2 t^2 + \mu t} (\because Z \text{ moments substitutable})$$

$$F_Z(z) = P(Z \leq z) = \Phi(z) = \frac{1}{2} + \frac{1}{2} \text{erf}(\frac{z}{\sqrt{2}}) \mid P(X \in [\mu - t\sigma, \mu + \sigma]) = \text{erf}(\frac{t}{\sqrt{2}})$$

Limits:  $\lim_{n \rightarrow \infty} \text{Binom}(N, p) \equiv \text{N}(Np, Np(1-p)), \lim_{n \rightarrow \infty} \text{Po}(X) \equiv \text{N}(\lambda, \lambda), \lim_{k \rightarrow \infty} \chi^2(k) \equiv \text{N}(2k, 2k), \lim_{v \rightarrow \infty} t_v \equiv Z$

## Lognormal (Galton) continuous RV X ~ Galton(μ, σ²)

$$f_X(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}} \quad F_X(x) = \int_{-\infty}^x f_X dx = \frac{1}{2} + \frac{1}{2} \text{erf}(\frac{x-\mu}{\sqrt{2\sigma^2}}) \quad E[X] = \text{Med} = \text{Mode} = \mu, E[X^2] = \mu^2 + \sigma^2$$

$$\text{Student's T continuous RV X} \sim t_n = \frac{Z}{\sqrt{V/n}}, V \sim Z_1^2 + \dots + Z_n^2 \text{ IID } \chi^2(n) \text{ each independent from } Z$$

$$f_X(x) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{\pi n \Gamma(\frac{n}{2})}} (1 + \frac{x^2}{n})^{-\frac{n+1}{2}} = \frac{1}{\sqrt{n} B(\frac{1}{2}, \frac{n}{2})} (1 + \frac{x^2}{n})^{-\frac{n+1}{2}} \quad \text{Mean} = \text{Mode} = 0. \quad \forall n \in \mathbb{N}^+ \forall \alpha \in [0, \frac{1}{2}] t_{\alpha/2, n-2} > z_{\alpha/2}$$

$$\forall k \in [0, n) (E[X^k] = \begin{cases} 0 & \forall k \equiv 1 \pmod 2 \\ \frac{1}{\sqrt{\pi \Gamma(\frac{n}{2})}} \Gamma(\frac{k+1}{2}) \Gamma(\frac{n-k}{2})^{\frac{1}{2}} & \forall n \equiv 0 \pmod 2 \end{cases} \Rightarrow \text{Var}(X) = \frac{1}{n-2} \prod_{i=1}^{k/2} \frac{2i-1}{n-2i} \quad \forall n \equiv 1 \pmod 2 \\ \text{N.A} \quad \forall n \leq 1 \end{cases}$$

## Cauchy continuous RV X ~ Cauchy(θ, γ) ≡ t₁

$$f_X(x) = \frac{1}{\pi \gamma (1 + \frac{x-\theta}{\gamma})^2} \quad F_X(x) = \frac{1}{\pi} \tan^{-1}(\frac{x-\theta}{\gamma}) + \frac{1}{2}. \quad \forall k \in \mathbb{N}^+ (E[X^k] \text{ N.A}). \quad \text{Median} = \text{Mode} = \theta$$

## Euler Integrals & Gamma Functions

Euler Integral 1st kind (Beta Function):  $\forall \alpha, \beta > 0 (B(\alpha, \beta) = (\frac{1}{\alpha} + \frac{1}{\beta}) \frac{1}{\alpha + \beta})$

Euler Integral 2nd kind (Gamma Function):  $\Gamma(z) = \int_0^{\infty} e^{-t} t^{z-1} dt$  | Upper-incomplete  $\Gamma(z, x) = \int_x^{\infty} e^{-t} t^{z-1} dt$

$\therefore \Gamma(\alpha)\Gamma(\beta) = \Gamma(\alpha + \beta)B(\alpha, \beta).$   $\forall z \in \mathbb{R} (\Gamma(z+1) = z\Gamma(z)).$   $\forall z \in \mathbb{Z} (\Gamma(z) = (z-1)!).$   $\Gamma(\frac{1}{2}) = \sqrt{\pi}$

## Gamma continuous RV X ~ Gamma(α, λ)

$$f_X(x) = \frac{\lambda^{\alpha} e^{-\lambda x} x^{\alpha-1}}{\Gamma(\alpha)}, x \geq 0. \quad \forall k \in \mathbb{N}^+ (E[X^k] = \frac{\Gamma(\alpha+k)}{\lambda^k}) \Rightarrow \text{Var}(X) = \frac{\alpha}{\lambda^2}, \text{Mode} = \frac{\alpha-1}{\lambda} \quad \forall \alpha \geq 1$$

## $\chi^2$ continuous RV X ~ $\chi^2(k) \equiv \text{Gamma}(\frac{k}{2}, \frac{1}{2}) = Z_1^2 + \dots + Z_k^2$ IID $\text{N}(0, 1)$ 's

$$f_X(x) = \frac{e^{-x/2} x^{k/2-1}}{2^{k/2} \Gamma(k/2)}, x \geq 0. \quad E[X] = k, E[X^2] = k(k+2), E[X^3] = k(k+2)(k+4). \quad \forall m > -\frac{k}{2} (E[X^m] = \frac{2^m \Gamma(\frac{k+m}{2})}{\Gamma(\frac{k}{2})}).$$

Var(X) =  $2k$ , Med ≈  $k(1 - \frac{2}{9k})^3$ , Mode = max{2k-2, 0}

## Exponential continuous RV X ~ Exp(λ) ≡ Gamma(1, λ)

$f_X(x) = \lambda e^{-\lambda x}, x \geq 0.$   $F_X(x) = 1 - e^{-\lambda x}$  memoryless like geometric i.e.,  $P(X > s+t) = P(X > s)P(X > t)$

$$E[X] = \frac{1}{\lambda}, E[X^2] = \frac{2}{\lambda^2}, E[X^3] = \frac{6}{\lambda^3}. \quad \forall k \in \mathbb{N}^+ (E[X^k] = \frac{k!}{\lambda^k}). \quad \text{Var}(X) = \frac{1}{\lambda^2}, \text{Median} = \frac{\ln(2)}{\lambda}, \text{Mode} = 0$$

## Pareto continuous RV X ~ Pareto(α, λ) = $\alpha e^{-\lambda X}, Y \sim \text{Exp}(\lambda)$

$$f_X(x) = \frac{\lambda^{\alpha} \lambda}{x^{\alpha+1}}. \quad F_X(x) = 1 - (\frac{\alpha}{x})^{\lambda}. \quad \forall k \in \mathbb{N} (E[X^k] = \begin{cases} \frac{\lambda^{\alpha} \lambda^k}{x^{\alpha+k}} & \forall \lambda > k \\ \text{N.A} & \forall \lambda \leq k \end{cases} \quad \alpha^k m(\text{Exp}(\lambda)), \text{Var}(X) = \begin{cases} \frac{\lambda \alpha^2}{(\lambda-1)(\lambda-2)} & \forall \lambda > 2 \\ \infty & \forall \lambda \leq 2 \end{cases}$$

## Laplace continuous RV X ~ Laplace(λ₁, λ₂) = $\text{Exp}(\lambda_1) - \text{Exp}(\lambda_2)$

$$f_X(x) = \frac{1}{2} e^{-\lambda |x|}. \quad F_X(x) = \begin{cases} 0.5 e^{-\lambda(\mu-x)} & \forall x \leq \mu \\ 1 - 0.5 e^{-\lambda(x-\mu)} & \forall x \geq \mu \end{cases} \quad E[X] = \text{Med} = \text{Mode} = \mu^2, E[X^2] = \mu^2 + \frac{2}{\lambda^2}, \text{Var}(X) = \frac{2}{\lambda^2}$$

## C5: Population (Non-random) Variables

**Definition**  
 Population variable: Oft-unknown, fixed (non-random) finite value estimated by random sample, summarised by parameters, visualised by histogram  
 Random draw w/o replacement has invariant distribution, so only violate independence in IID assumption.

## C6: Statistical Model for $\mu, \sigma^2$

Let  $x_1, \dots, x_n$  be realisations of IID RV  $X_1, \dots, X_n$  with replacement from same population with expectation  $\mu$ , variance  $\sigma^2$ .  $\therefore \bar{X} \sim \text{Dist}(\mu, \frac{\sigma^2}{n})$

## C7: Standard Error SE=SD(Estimator)

SE = SD(estimate  $\hat{\theta}$ ) fixed. Closure under scaling

- $\exists 4$  common distributions: 1.  $X \sim N(\mu, \sigma^2) \Leftrightarrow aX \sim N(a\mu, a^2\sigma^2)$   
 2.  $X \sim \text{Galton}(\mu, \sigma^2) \Leftrightarrow Y = \ln(X) \sim N(\mu, \sigma^2) \Leftrightarrow \ln(aX) = Y + \ln(a) \sim N(\mu + \ln(a), \sigma^2)$   
 3.  $X \sim \text{Gamma}(\alpha, \lambda) \Leftrightarrow \beta X \sim \text{Gamma}(\alpha, \frac{\lambda}{\beta})$  | 4.  $X \sim \text{Exp}(\lambda) \Leftrightarrow \beta X \sim \text{Exp}(\frac{\lambda}{\beta})$

Non-eg:  $X \sim \text{Binom}(N, p) \Rightarrow aX$  may have non-ints, so non-binomial.

## C8: Statistical Model for Population Proportion p

Let  $x_1, \dots, x_n$  be realisations of IID Bernoulli( $p$ ) indicator RV  $X_1, \dots, X_n$  with replacement from same population with population proportion  $p$ .  $\therefore$  Random prop  $\hat{p} = \frac{s_n}{n} = \frac{\sum x_i}{n} \sim \text{Binomial}(n, p) \Rightarrow E[\hat{p}] = p$  unbiased,  $\text{Var}(\hat{p}) = \frac{p(1-p)}{n}$ .

ST2132 takes SE = SD( $\hat{p}$ ) =  $\sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$  with negative bias  $-\frac{p(1-p)}{n^2}$  i.e.  $p \approx \hat{p} \pm \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$ .

If w/o replacement, actual SE  $\approx \sqrt{\frac{N-n}{N-1}} \sqrt{\frac{\hat{p}(1-\hat{p})}{n}}$

## C9: Bias

f convex-downward  $\Leftrightarrow E[f(X)] \geq f(E[X])$

Bias( $\hat{\theta}$ ) =  $E[\hat{\theta}] - \theta$ , MSE( $\hat{\theta}$ ) =  $E[(\hat{\theta} - \theta)^2] = \text{Var}(\hat{\theta} - \theta) + \text{Bias}(\hat{\theta})^2 = \text{SE}^2 + b^2$

$\hat{\theta}$  consistent  $\Leftrightarrow \lim_{n \rightarrow \infty} \text{Bias}(\hat{\theta}) = 0$ , else intolerable for inferential statistics.

## C10: Confidence Interval

$z_{\alpha/2} = \sqrt{2}\text{erf}^{-1}(1 - \alpha)$ , significance level alpha = 1 - confidence level

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

**LU MINGYUAN [Student ID: A0269854J]**  
**ST2132 Mathematical Statistics**