1 Basic Properties

1. $E(X) = \sum xp(x)$
2. $Var(X) = \sum (x-\mu)^2 f(x)$
3. X is around $E(X)$, give or take $SD(X)$
4. $E(aX + bY) = aE(X) + bE(Y)$
5. $Var(aX + bY) = a^2 Var(X) + b^2 Var(Y)$
6. $Var(X) = E(X^2) - [E(X)]^2$
7. $Cov(X_1, X_2) = E(X_1 X_2) - E(X_1)E(X_2)$
8. $P(AB) = P(A)P(B)$ if A and B independent
9. RV is centered when $E(X) = 0$, and any RV can be centered via $Y = X - E(X)$, with SD and variance unaffected
10. In $X = \mu + \epsilon$, μ is the unknown constant of interest, and ϵ represents random measurement error.
11. if X, Y are independent:
    (a) $M_{X+Y}(t) = M_X(t) M_Y(t)$
    (b) $E(XY) = E(X)E(Y)$, converse is true if X and Y are bivariate normal, extends to multivariate normal

2 Approximations

2.1 Law of Large Numbers

Let $X_1, X_2, ..., X_n$ be IID, with expectation μ and variance σ^2. $\overline{X_n} = \frac{1}{n}\sum_{i=1}^n X_i \xrightarrow[n]{\infty} \mu$. Let $x_1, x_2, ..., x_n$ be realisations of the random variable $X_1, X_2, ..., X_n$, then $\overline{x_n} = \frac{1}{n}\sum_{i=1}^n x_n \xrightarrow[n]{\infty} \mu$

2.2 Central Limit Theorem

Let $S_n = \sum_{i=1}^n X_i$ where $X_1, X_2, ..., X_n$ IID. $\frac{S_n - n\mu}{\sqrt{n}\sigma} \xrightarrow[n]{\infty} \mathcal{N}(0,1)$

3 Distributions

3.1 Poisson(λ)

\$Pr(X=k) = {$\frac{}{\lambda^{ke}}$-λ}{k!}, k = 0,1,2,...\$
$E(X) = Var(X) = \lambda$

3.2 Normal $X \sim \mathcal{N}(\mu, \sigma^2)$

$f(x) = \frac{1}{\sqrt{2\pi}\sigma} exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right), -\infty < x < \infty$

1. When $\mu = 0$, $f(x)$ is an even function, and $E(X^k) = 0$ where k is odd
2. $Y = \frac{X - E(X)}{SD(X)}$ is the standard normal

3.3 Gamma Γ

$g(t) = \frac{\lambda^\alpha}{\Gamma(\alpha)} t^{\alpha-1} e^{-\lambda t}, t \geq 0$
\$$\mu_1 = \alpha\frac{}{\lambda, \mu_2 = \frac{\alpha(\alpha+1)}{\lambda^2}}$

3.4 χ^2 Distribution

Let $\mathcal{Z} \sim \mathcal{N}(0,1), \mathcal{U} = \mathcal{Z}^2$ has a χ^2 distribution with 1 d.f.
$f_{\mathcal{U}}(u) = \frac{1}{\sqrt{2\pi}} u^{-\frac{1}{2}} e^{-\frac{u}{2}}, u \geq 0$
$\chi_1^2 \sim \Gamma(\alpha = \frac{1}{2}, \lambda = \frac{1}{2})$
Let $U_1, U_2, ..., U_n$ be χ_1^2 IID, then $V = \sum_{i=1}^n U_i$ is χ_n^2 with n degree freedom, $V \sim \Gamma(\alpha = \frac{n}{2}, \lambda = \frac{1}{2})$
$E(\chi_n^2) = n, Var(\chi_n^2) = 2n$
$M(t) = (1-2t)^{-\frac{n}{2}}$

3.5 t-distribution

Let $\mathcal{Z} \sim \mathcal{N}(0,1)$, $\mathcal{U}_n \sim \chi_n^2$ be independent, $t_n = \frac{\mathcal{Z}}{\sqrt{\mathcal{U}_n/n}}$ has a t-distribution with n d.f.

$f(t) = \frac{\Gamma([(n+1)/2])}{\sqrt{n\pi}\Gamma(n/2)}\left(1 + \frac{t^2}{n}\right)^{-\frac{n+1}{2}}$

1. t is symmetric about 0
2. $t_n \xrightarrow[n]{\infty} \mathcal{Z}$

3.6 F-distribution

Let $U \sim \chi_m^2, V \sim \chi_n^2$ be independent, $W = \frac{U/m}{V/n}$ has an F distribution with (m,n) d.f.
If $X \sim t_n$, $X^2 = \frac{\mathcal{Z}/1}{\mathcal{U}_n/n}$ is an F distribution with (1,n) d.f, with $w \geq 0$:
$f(w) = \frac{\Gamma([(n+1)/2])}{\Gamma(m/2)\Gamma(n/2)} \frac{m}{n}^{\frac{m}{2}} w^{\frac{m}{2}-1}\left(1 + \frac{m}{n}w\right)^{-\frac{m+n}{2}}$
For $n > 2$, $E(W) = \frac{n}{n-2}$

4 Sampling

Let $X_1, X_2, ..., X_n$ be IID $\mathcal{N}(\mu, \sigma^2)$.
sample mean, $\overline{X} = \frac{1}{n}\sum_{i=1}^n X_i$
sample variance, $S^2 = \frac{1}{n-1}\sum_{i=1}^n \left(X_i - \overline{X}\right)^2$

4.1 Properties of $\overline{X}$ and S^2

1. $\overline{X}$ and S^2 are independent
2. $\overline{X} \sim \mathcal{N}(\mu, \frac{\sigma^2}{n})$
3. $\frac{(n-1)S^2}{\sigma^2} \sim \chi_{n-1}^2$
4. $\frac{\overline{X} - \mu}{S/\sqrt{n}} \sim t_{n-1}$

4.2 Survey Sampling

In population of size N, we are interested in a variable x. The ith individual has fixed value x_i.
mean of population $= \mu = \frac{1}{N}\sum_{i=1}^N x_i$
total of population $= \tau = \sum_{i=1}^N x_i = \mu N$
SD of population $= \sigma$
$\sigma^2 = \sum_{i=1}^N (x_i - \mu)^2 \frac{1}{N}\sum_{i=1}^n x_i^2 - \mu^2$

4.2.1 Dichotomous case

Population are members with value 0 or 1. Let p be the proportion of members with value 1. $\mu = p, \sigma^2 = p(1-p)$

4.3 Simple Random Sampling (SRS)

Assume n random draws are made without replacement. (Not SRS, will be corrected for later).

4.3.1 Lemma A

The draws X_i have the same distribution, and denote $\xi_1, \xi_2, ...\xi_n$ as values assumed by the population, and let the number of members with value ξ_j be n_j
$P(X_i = \xi_j) = \frac{n_j}{N}$
$E(X_i) = \mu, Var(x_i) = \sigma^2$

4.3.2 Lemma B

For $i \neq j$, $Cov(X_i, X_j) = -\frac{\sigma^2}{N-1}$
We use sample mean $\overline{X}$ to estimate μ:
$E(\overline{X}) = \mu$ from Lemma A, and
$Var(\overline{X}) = \frac{\sigma^2}{n}\left(\frac{N-n}{N-1}\right)$ from Lemma B, where $\frac{N-n}{N-1}$ is the finite population correction factor.
In 0-1 population, let $\hat{p}$ be proportion of 1s in the sample:
$E(\hat{p}) = p, SD(\hat{p}) = \sqrt{\frac{p(1-p)}{n}\frac{N-n}{N-1}}$

4.3.3 Estimation Problem

Let $X_1, X_2, ..., X_n$ be random draws with replacement. Then $\overline{X}$ is an estimator of μ. and the observed value of $\overline{X}$, $\overline{x}$ is an estimate of μ.

4.3.4 Standard Error (SE)

Since $E(\overline{X}) = \mu$, the estimator is unbiased.
The error in a particular estimate $\overline{X}$ is unknown, but on average its size is about $SD(\overline{x}) = \frac{\sigma}{\sqrt{n}}$
Standard error of an $\overline{X}$ is defined to be $SD(\overline{X})$
An unbiased estimator for σ^2 is $s^2 = \frac{1}{n-1}\sum_{i=1}^n (X_i - \overline{X})^2$

param	est	SE	Est. SE
μ	$\overline{X}$	$\frac{\sigma}{\sqrt{n}}$	$\frac{s}{\sqrt{n}}$
p	$\hat{p}$	$\sqrt{\frac{p(1-p)}{n}}$	$\sqrt{\frac{\hat{p}(1-\hat{p})}{n-1}}$

4.3.5 Without Replacement

SE is multiplied by $\frac{N-n}{N-1}$, because s^2 is biased for σ^2: $E(\frac{N-1}{N}s^2) = \sigma^2$, but N is normally large.

4.3.6 Confidence Interval

An approximate $1 - \alpha$ CI for μ is $(\overline{x} - z_{\alpha/2}\frac{s}{\sqrt{n}}, \overline{x} + z_{\alpha/2}\frac{s}{\sqrt{n}})$

4.4 Measurement Error

Let $x_1, x_2, ..., x_n$ be independent measurements of unknown constant μ. $X_i = \mu + \epsilon_i$.
The errors are IID with expectation 0 , and variance σ^2. $x_i = \mu + e_i$, where x_i and e_i are realisations of the RV. Then $\overline{x}$ is an estimate of μ, with SE $\frac{\sigma}{\sqrt{n}}$.

4.4.1 Biased Measurements

Let $X = \mu + \epsilon$, where $E(\epsilon) = 0$, $Var(\epsilon) = \sigma^2$
Suppose X is used to measure an unknown constant a, $a \neq \mu$. $X = a + (\mu - a) + \epsilon$, where $\mu - a$ is the bias.
Mean square error (MSE) is $E((X-a)^2) = \sigma^2 + (\mu - a)^2$
with n IID measurements, $\overline{x} = \mu + \overline{\epsilon}$
$E((x-a)^2) = \frac{\sigma^2}{n} + (\mu - a)^2$
MSE = SE^2 + bias^2, hence $\sqrt{\text{MSE}}$ is a good measure of the accuracy of the estimate $\overline{x}$ of a.

4.5 Estimation of a Ratio

Consider a population of N members, and two characteristics are recorded: $(X_1, Y_1), (X_2, Y_2), ..., (X_n, Y_n)$, $r = \frac{\mu_y}{\mu_x}$.
An obvious estimator of r is $R = \frac{\overline{Y}}{\overline{X}}$
$Cov(\overline{X}, \overline{Y}) = \frac{\sigma_{xy}}{n}$, where
$\sigma_{xy} := \frac{1}{N}\sum_{i=1}^N (x_i - \mu_x)(x_i - \mu_y)$ is the population covariance.

4.5.1 Properties

With SRS, the approx variance of $R = \overline{Y}/\overline{X}$ is

$$Var(R) \approx \frac{1}{\mu_x^2}\left(r^2\sigma_{\overline{X}}^2 + \sigma_{\overline{Y}}^2 - 2r\sigma_{\overline{XY}}\right)$$
$$= \frac{1}{n}\frac{N-n}{N-1}\frac{1}{\mu_x^2}\left(r^2\sigma_{\overline{X}}^2 + \sigma_{\overline{Y}}^2 - 2r\sigma_{\overline{XY}}\right)$$

Population coefficient $\rho = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$

$E(R) \approx r + \frac{1}{n}\left(\frac{N-n}{N-1}\right)\frac{1}{\mu_x^2}\left(r\sigma_x^2 - \rho\sigma_x\sigma_y\right)$

$s_{xy} = \frac{1}{n-1}\sum_{i=1}^n\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)$

4.5.2 Ratio Estimates

$\overline{Y}_R = \frac{\mu_x}{\overline{X}}\overline{Y} = \mu_x R$

$Var(\overline{Y}_R) \approx \frac{1}{n}\frac{N-n}{N-1}(r^2\sigma_x^2 + \sigma_y^2 - 2r\rho\sigma_x\sigma_y)$

$E(\overline{Y}_R) - \mu_y \approx \frac{1}{n}\frac{N-n}{N-1}\frac{1}{\mu_x}\left(r\sigma_x^2 - \rho\sigma_x\sigma_y\right)$

The bias is of order $\frac{1}{n}$, small compared to its standard error.

$\overline{Y}_R$ is better than $\overline{Y}$, having smaller variance, when $\rho > \frac{1}{2}\left(\frac{C_x}{C_y}\right)$, where $C_i = \sigma_i/\mu_i$

Variance of $\overline{Y}_R$ can be estimated by

$s_{\overline{Y}_R}^2 = \frac{1}{n}\frac{N-n}{N-1}\left(R^2 s_x^2 + s_y^2 - 2Rs_{xy}\right)$

An approximate $1-\alpha$ C.I. for μ_y is $\overline{Y}_R \pm z_\alpha/2 s_{\overline{Y}_R}$

5 Estimation

Let $X_1, X_2, ..., X_n$ be IID random variables with density $f(x|\theta)$, where $\theta \in \mathcal{R}^P$ is an unknown constant. Realisations $x_1, x_2, ..., x_n$ will be used to estimate θ, the estimate a realisation of RV $\hat{\theta}$. The bias and SE are:

bias $= E(\hat{\theta}) - \theta, SE = SD(\hat{\theta})$

5.1 Moments

Let $X_1, X_2, ..., X_n$ be IID with the same distribution as X.

$\hat{\mu}_k = \frac{1}{n}\sum_{i=1}^n X_i^k$ is an estimator of μ_k, where μ_k is the kth moment. An estimate is also denoted $\hat{\mu}_k$.

5.2 Method of Moments

To estimate θ, express it as a function of moments $g(\hat{\mu}_1, \hat{\mu}_2, ...)$

The bias and SE in an estimate, still depends on the unknown value of the constant. Suppose 1.67 and 0.38 are estimates of λ and α. Data is generated from $\Gamma(1.67, 0.38)$, and the MOM estimators are written as $\widehat{1.67}$ and $\widehat{0.38}$. Because the sample size is large, $(\hat{\lambda} - \lambda, \hat{\alpha} - \alpha) \approx (\widehat{1.67} - 1.67, \widehat{0.38} - 0.38)$

Monte Carlo is used to generate many realisations of $\widehat{1.67}$ via the $\Gamma(1.67, 0.38)$ distribution. With 10,000 realisations,

$bias(1.67) = E_{1.67,0.38}(\widehat{1.67} - 1.67) \approx 0.09$

$SE(1.67) = SD_{1.67,0.38}(\widehat{0.38}) \approx 0.35$

and λ is estimated as 1.58 ± 0.35

$\overline{X} \xrightarrow[n]{\infty} \alpha/\lambda, \hat{\sigma}^2 \xrightarrow[n]{\infty} \alpha/\lambda^2$, MOM estimators are consistent (asymptotically unbiased).