S. A. Klugman, H. H. Panjer, and G. E. Willmot,
Chapter 12: "Review of Mathematical Statistics,"
in *Loss Models: From Data to Decisions*, Third Edition, pp. 315–30.

OUTLINE

I. POINT ESTIMATION

A. Introduction

1. Definitions
   a. **Estimate** – "specific value obtained when applying an estimation procedure to a set of numbers"
   b. **Estimator** – random variable or random function that comprises a rule or formula producing an estimate
   c. **Model error** – error resulting from the selection of the wrong model
   d. **Sampling error** – error resulting from making inferences about the total population based on a sample from that population
   e. **Sampling frame error** – error resulting from trying to draw inferences about a population that differs from the one sampled
2. Modeling
   a. Need to select both parameters and type of a model used, usually in that order
   b. Potential error in parameter estimates expressed via an interval estimate
   c. Cannot determine the amount of error in a specific estimate but rather that inherent in the estimation procedure used
   d. Random future experience does not invalidate the actuarial assumptions assumed
3. Examples of errors
   a. Model error – choosing a Poisson when the actual distribution is a negative binomial
   b. Sampling frame error – applying conclusions regarding policies from independent agents to those sold through the mail

B. Measures of Quality

1. Definitions
   a. **Asymptotically unbiased** – characteristic of an estimator when its expected value approaches the true parameter value as the sample size approaches infinity

$$\lim_{n\to\infty} E[\hat{\theta}_n \mid \theta] = \theta$$

   b. **Bias** – difference between the expected value of the estimator and the true parameter value

$$\text{Bias}_{\hat{\theta}}(\theta) = E[\hat{\theta} \mid \theta] - \theta$$

c. **Consistent** a.k.a. **weakly consistent** – characteristic of an estimator if for all $\delta > 0$ and any θ, we have

$$\lim_{n\to\infty} P\left(|\hat{\theta}_n - \theta| > \delta\right) = 0$$

d. **Mean squared error (MSE) of an estimator**

$$MSE_{\hat{\theta}}(\theta) = E\left[(\hat{\theta} - \theta)^2 \mid \theta\right] = Var\left(\hat{\theta} \mid \theta\right) + [Bias_{\hat{\theta}}(\theta)]^2$$

e. **Relative efficiency** – ratio of the MSEs for two estimators

f. **Unbiased estimator** – characteristic of estimator $\hat{\theta}$ if $E\left[\hat{\theta} \mid \theta\right] = \theta$ for all θ

g. **Uniformly minimum variance unbiased estimator (UMVUE)** – unbiased estimator such that "for any true value of θ there is no other unbiased estimator that has a smaller variance"

2. Comments

a. Sufficient (but not necessary) condition for an estimator to be weakly consistent is that it be asymptotically unbiased and $Var\left(\hat{\theta}_n\right) \to 0$

b. Maximum observation from a uniform distribution on $(0, \theta)$

1) Is an asymptotically unbiased estimator of θ

2) Is a consistent estimator of θ

II. INTERVAL ESTIMATION

A. Definitions

1. **Interval estimator** – range of possible numbers, each likely to be the true value
2. **Level of confidence** – probability parameter θ will be in the interval (L, U)
3. **Point estimator** – "single value that represents our best attempt to determine the value of the unknown population quantity"
4. **100(1 – α)% confidence interval** for a parameter θ – interval estimator comprised of a pair of values L and U computed from a random sample such that $P(L \le \theta \le U) \ge (1 - \alpha)$ for all θ

B. Construction of Confidence Intervals

1. Based on a t distribution

$$L = \bar{X} - t_{\alpha/2,n-1}\, s/\sqrt{n} \qquad U = \bar{X} + t_{\alpha/2,n-1}\, s/\sqrt{n}, \text{ where}$$

$$s = \sqrt{\left[\sum_{j=1}^{n}(X_j - \bar{X})^2\right] / (n-1)}$$

$t_{\alpha/2,n-1}$ - $100(1 - \alpha/2)$th percent of the t distribution with $(n - 1)$ degrees of freedom

PAST CAS AND SoA EXAMINATION QUESTIONS

A. Measures of Quality

A23. According to Klugman et al., which of the following must be true of a consistent estimator?

1. It is unbiased.
2. For any small quantity ε, the probability that the absolute value of the deviation of the estimator from the true parameter value is less than ε tends to 1 as the number of observations tends to infinity.
3. It has minimal variance.

A. 1 B. 2 C. 3 D. 2,3 E. 1,2,3 (96S–4B–12–1)

A24. You are given the following:

i) The expectation of a given estimator is .50.
ii) The variance of this estimator is 1.00.
iii) The bias of this estimator is .50.

Determine the mean square error of this estimator.

A. .75 B. 1.00 C. 1.25 D. 1.50 E. 1.75 (96F–4B–21–2)

A25. $MSE_{\hat{\theta}}(\theta) = Var(\hat{\theta}) + [Bias_{\hat{\theta}}(\theta)]^2$ (00S–4–MC)

A26. If $\hat{\theta}_n$ is asymptotically unbiased and $\lim_{n\to\infty} Var(\hat{\theta}_n) = 0$, then $\hat{\theta}_n$ is weakly consistent. (00S–4–MC)

A27. You are given two independent estimates of an unknown quantity μ:

i) Estimate A: $E[\mu_A] = 1{,}000$ and $\sigma(\mu_A) = 400$
ii) Estimate B: $E[\mu_B] = 1{,}200$ and $\sigma(\mu_B) = 200$

Estimate C is a weighted average of the two estimates A and B, such that:

$$\mu_C = w\mu_A + (1 - w)\mu_B$$

Determine the value of w that minimizes $\sigma(\mu_C)$.

A. 0 B. 1/5 C. 1/4 D. 1/3 E. 1/2 (00S–4–18)

A28. You are given:

x	0	1	2	3
Pr[X = x]	.5	.3	.1	.1

The method of moments is used to estimate the population mean, μ, and variance, σ^2, by $\bar{X}$ and $S_n^2 = \frac{\Sigma (X_i - \bar{X})^2}{n}$, respectively. Calculate the bias of S_n^2, when n = 4.

A. –.72 B. –.49 C. –.24 D. –.08 E. 0 (02F–4–31) (Sample–C–49)

Solutions are based on pp. 316–24 plus the pages cited.

A23. 1. F – This is not required.

2. T

3. F – This is not required.

Answer: B

A24. $MSE_{\hat{\theta}}(\theta) = Var(\hat{\theta}) + [b_{\hat{\theta}}(\theta)]^2 = 1.00 + (.50)^2 = 1.25$

Answer: C

A25. T.

A26. T.

A27. $Var(\mu_C) = w^2 Var(\mu_A) + (1 - w)^2 Var(\mu_B) = w^2(400)^2 + (1 - w)^2(200)^2$

$\partial Var(Z)/\partial w = 2w(400)^2 + (2w - 2)(200)^2 = 0 \qquad w = .2$

Answer: B

A28. $E\left[S_n^2\right] = (1/n)\, E\left[\sum_{j=1}^{n} (X_j - \bar{X})^2\right] = \left(1/n\right)\left(E\left[\sum_{j=1}^{n} X_j^2\right] - E[n\bar{X}^2]\right)$

$E\left[S_n^2\right] = (\sigma^2 + \mu^2) - (1/n^2)\sum_{i=1}^{n}\sum_{j=1}^{n} E[X_i X_j]$

$E\left[S_n^2\right] = (\sigma^2 + \mu^2) - [1/n^2][n(n - 1)\mu^2 + n(\sigma^2 + \mu^2)] = \dfrac{(n - 1)\sigma^2}{n}$

$\mu = (0)(.5) + (1)(.3) + (2)(.1) + (3)(.1) = .8$

$\sigma^2 = (0 - .8)^2(.5) + (1 - .8)^2(.3) + (2 - .8)^2(.1) + (3 - .8)^2(.1) = .96$

$E\left[S_n^2\right] = \dfrac{(n - 1)\sigma^2}{n}$

$Bias\left(S_n^2\right) = E\left[S_n^2\right] - \sigma^2 = \dfrac{(n - 1)\sigma^2}{n} - \sigma^2 = -\sigma^2/n = -.96/4 = -.24$

Answer: C