WMS: 1

I. Aims of Statistics:
  A. Probability is the study of relative proportions of random outcomes based on structure of generating process.
  B. Statistics is the inverse problem:
    1. Observe data generated by process
    2. Infer process.

WMS: 8.1

II. Estimation
  A. Aim
    1. Want to estimate some number θ, called a *parameter*.
      a. Fraction of population supporting a candidate
      b. Population Average effect of some cholesterol–lowering medication.
      c. Mass of an electron
    2. Rule that gives estimate is called an estimator.
    3. Want it based on some data.

WMS: 8.2, 8.4

  B. Preliminaries: What makes a good estimator
    1. Quantify what happens if you make a wrong decision
      a. Suppose that you pay a penalty $L(a,\theta)$ if your guess is a when the truth is θ
        i. Penalty is called *Loss function*.
        ii. Most typically, $L(a,\theta)=(a-\theta)^2$: Squared error loss.
  C. Typically, want rule that depends on data, $\delta(\boldsymbol{X})$
    1. Example
      a. If $X\sim \mathcal{B}\text{in}(n,\theta)$, $\delta(X)$ might be X/n

      b. If $\boldsymbol{X}$ is a sample from a population with expectation θ, then $\delta(\boldsymbol{X})$ might be $\bar{X}$
      c. If $\boldsymbol{X}$ is a sample from a population with median θ, then $\delta(\boldsymbol{X})$ might be sample median
    2. Select estimator before wee see data.
      a. Consider average of loss function $R(\delta,\theta)=\mathrm{E}\left[L(\delta(\boldsymbol{X}),\theta)\right]$.
      b. R is called *risk function*.
        i. Risk function for squared error loss is called *mean squared error (MSE)*.
      c. Review expectation
      d. Let $\mu=\mathrm{E}\left[\delta(\boldsymbol{X})\right]$
      e. $R(\delta,\theta)$ is variance plus bias squared.

$$
\begin{aligned}
&=\mathrm{E}\left[(\delta(\boldsymbol{X})-\mu+\mu-\theta)^2\right]\\
&=\mathrm{E}\left[(\delta(\boldsymbol{X})-\mu)^2+2(\delta(\boldsymbol{X})-\mu)(\mu-\theta)+(\mu-\theta)^2\right]\\
&=\mathrm{E}\left[(\delta(\boldsymbol{X})-\mu)^2\right]+\mathrm{E}\left[2(\delta(\boldsymbol{X})-\mu)(\mu-\theta)\right]\\
&\quad+\mathrm{E}\left[(\mu-\theta)^2\right]\\
&=\mathrm{Var}\left[\delta(\boldsymbol{X})\right]+2(\mu-\theta)\mathrm{E}\left[(\delta(\boldsymbol{X})-\mu)\right]+(\mu-\theta)^2\\
&=\mathrm{Var}\left[\delta(\boldsymbol{X})\right]+(\mathrm{E}\left[\delta(\boldsymbol{X})\right]-\theta)^2
\end{aligned}
$$

      f. We can often make the second part 0:
        i. If $\mathrm{E}\left[\delta(\boldsymbol{X})\right]=\theta$, then $\delta(\boldsymbol{X})$ is called *unbiased*.
        ii. and $\mathrm{E}\left[\delta(\boldsymbol{X})\right]-\theta$ is called the *bias*.
      g. We will see that unbiasedness does not completely specify the best estimator
        i. Ex.: $X\sim\mathcal{N}(\theta,\sigma^2)$ with σ known.
        ii. $\delta(X)=aX+b$

        iii. $R(\delta,\theta)=(a\theta+b-\theta)^2+a^2\sigma^2=(a-1)^2\theta^2+2b(a-1)\theta+b^2+a^2\sigma^2$
        iv. If $a\neq 1$ maximum is $\infty \Rightarrow$ choose $a=1$, and risk is $b^2+a^2\sigma^2 \Rightarrow$ choose $b=0$
      h. “Best” (ie, minimax) estimator might allow some bias in return for smaller variance

WMS: 8.3

  D. Examples
    1. Estimate of the range of a uniform distribution from the range of a sample
      a. $X_1,\ldots,X_n\sim\mathcal{U}[\alpha,\beta]$, i.i.d..
      b. $X_{(1)},\ldots,X_{(n)}$ are ordered values: order statistics.
      c. $\delta(\boldsymbol{X})=X_{(n)}-X_{(1)}$
      d. Density of $X_{(n)}$ is

$$n\frac{1}{\beta-\alpha}\left(\frac{y-\alpha}{\beta-\alpha}\right)^{n-1}$$

      e. $\mathrm{E}\left[X_{(n)}\right]$ is

$$
\begin{aligned}
&=\int_\alpha^\beta yn\frac{1}{\beta-\alpha}\left(\frac{y-\alpha}{\beta-\alpha}\right)^{n-1}\,dy\\
&=\int_0^1(\alpha+z(\beta-\alpha))nz^{n-1}\,dz\\
&=n\alpha\int_0^1 z^{n-1}\,dz+n(\beta-\alpha)\int_0^1 z^n\,dz\\
&=\alpha+\frac{n}{n+1}(\beta-\alpha)
\end{aligned}
$$

      f. $\mathrm{E}\left[Y_{(1)}\right]=-(-\beta+\frac{n}{n+1}(-\alpha-(-\beta))=\beta-\frac{n}{n+1}(\beta-\alpha)$

      g. $\mathrm{E}\left[\text{Range}\right]=\alpha-\beta+\frac{2n}{n+1}(\beta-\alpha)=e_n(\beta-\alpha)$ for $e_n=\frac{n-1}{n+1}$.
      h. Bias is $\frac{-2}{n+1}(\beta-\alpha)$
        i. Almost unbiased.
          - $\lim_{n\to\infty}$ bias $=0$.
          - Called *asymptotically unbiased*.
      i. Hence to get unbiased estimator, use $\frac{n+1}{n-1}$Range.
        i. Let $v_n=\mathrm{Var}\left[X_{(n)}-X_{(1)}\right](\beta-\alpha)^{-2}$, which does not depend on α or β.
        ii. MSE of $\delta(\boldsymbol{X})=a\text{Range}(\boldsymbol{X})$ is $\{(ae_n-1)^2+a^2v_n\}(\beta-\alpha)^2$
        iii. Differentiating, setting the derivative to zero, and solving for a gives $a=e_n/(e_n^2+v_n)$.
        iv. Since $v_n>0$, minimizing a leaves $\delta(\boldsymbol{X})$ with a slight bias.
    2. Estimating variance
      a. $X_1,\ldots,X_n$ i.i.d., expectation μ, variance σ^2.
      b. $\bar{X}=\sum_{j=1}^n X_j/n$
      c. Then $\sum_{j=1}^n(X_j-\mu)^2$ is

$$=\sum_{j=1}^n(X_j-\bar{X})^2+\sum_{j=1}^n(\bar{X}-\mu)^2+2\sum_{j=1}^n(\bar{X}-\mu)(X_j-\bar{X})^2$$

$$=\sum_{j=1}^n(X_j-\bar{X})^2+n(\bar{X}-\mu)^2.$$

      d. $\sum_{j=1}^n(X_j-\bar{X})^2=\sum_{j=1}^n(X_j-\mu)^2-n(\mu-\bar{X})^2$
      e. $S^2=\sum_{j=1}^n(X_j-\bar{X})^2/(n-1)$

i. $= [\sum_{j=1}^{n}(X_j-\mu)^2 - n(\bar{X}-\mu)^2]/(n-1)$

f. $\mathrm{E}\left[S^2\right] = \dfrac{\sum_{i=1}^{n}\mathrm{E}\left[(X_i-\mu)^2\right]}{n-1} - n\dfrac{\mathrm{E}\left[(\bar{X}-\mu)^2\right]}{n-1}$.

$= n\sigma^2/(n-1) - \sigma^2/(n-1) = \sigma^2$

i. Makes no assumption about distribution of observations.

3. Estimating Standard Deviation
   a. $X_1,\ldots,X_n \sim \mathcal{N}[\mu,\sigma^2]$, i.i.d..
   b. $(n-1)S^2/\sigma^2 \sim \chi^2(n-1)$.
      i. χ_k^2 distribution is distribution of sum of squares of k independent $\mathcal{N}(0,1)$ random variables
   c. $\mathrm{E}\left[S\sqrt{n-1}/\sigma\right]$

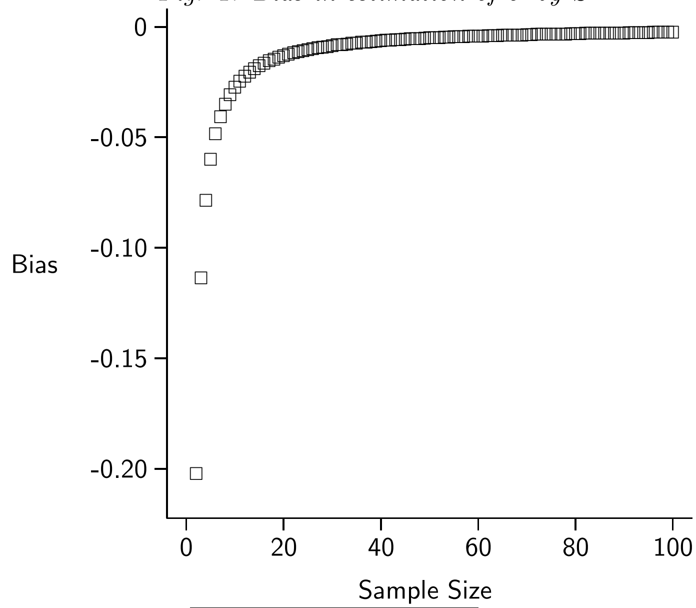
$$= \int_0^\infty \frac{\sqrt{x}\exp(-1/2x)x^{(n-1)/2-1}\,dx}{2^{(n-1)/2}\Gamma((n-1)/2)}$$

$$= \sqrt{2}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}\int_0^\infty \frac{\exp(-1/2x)x^{n/2-1}\,dx}{2^{n/2}\Gamma(n/2)}$$

$$= \sqrt{2}\frac{\Gamma(n/2)}{\Gamma((n-1)/2)}.$$

   d. $\mathrm{E}[S] = \sigma\dfrac{\sqrt{2}\Gamma(n/2)}{\Gamma((n-1)/2)\sqrt{n-1}}$
   e. Bias is $\sigma\left(\dfrac{\sqrt{2}\Gamma(n/2)}{\Gamma((n-1)/2)\sqrt{n-1}} - 1\right)$
   f. See Fig. 1.

Fig. 1: Bias in estimation of σ by S



Calculations for $\sqrt{\sum_{k=1}^{n}(X_k-\bar{X})^2/(n-1)}$ with normal observations

This page intentionally left blank.

This page intentionally left blank.