Emma Erdtmann and Daniel Detore

I pledge my honor that I have abided by the Stevens Honor System.

Problem 1: Verify the joint pmf of the Bernoulli sample on page 9 of Slides02 and show that it depends on the sample only through the sample mean.

Slides02, page 7

- Bernoulli population $X \sim \mathcal{B}(1, p)$ The cdf and pmf are

$$F_2(x, \theta) = \begin{cases} 0, & \text{if } x < 0, \\ 1 - p, & \text{if } 0 \le x < 1, \\ 1, & \text{if } x \ge 1, \end{cases} \qquad f_2(x, \theta) = \begin{cases} 1 - p, & \text{for } x = 0, \\ p, & \text{for } x = 1, \end{cases}$$

respectively, and the one-dimensional parameter $\theta = p$.

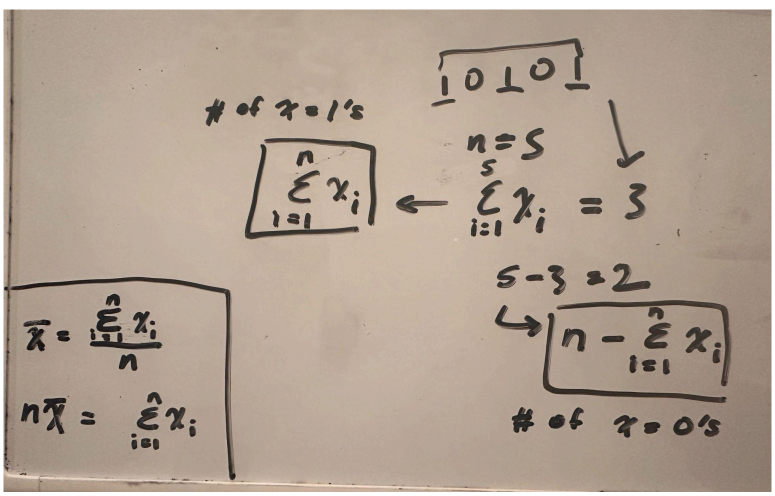
Slides02, page 9

If $X \sim \mathcal{B}(1, p)$, then $(X_1, \ldots, X_n)$ is has joint pmf

$$\prod_{i=1}^{n} f_2(x_i, \theta) = p^{\sum_{i=1}^{n} x_i}(1 - p)^{n - \sum_{i=1}^{n} x_i},$$

where $\theta = p$ and $x_i = 0, 1$, $i = 1, \ldots, n$.

Step 1 & 2: Relating the sample mean to the summations, figuring out the summations for the number of times we have x = 0 and x = 1



Step 3 & 4: Finding the product as it relates to the number of x = 0's and x = 1's, then substituting that in with n*sample mean

$$\prod_{i=1}^{n} f_2(x_i, \theta) = p^{\#\text{ of } x=1\text{'s}} \cdot (1-p)^{\#\text{ of } x=0\text{'s}}$$

$$\hookrightarrow = \begin{cases} 1-p, & x=0 \\ p, & x=1 \end{cases}$$

$$= p^{\sum_{i=1}^{n} x_i} \cdot (1-p)^{n-\sum_{i=1}^{n} x_i}$$

$$\boxed{= p^{n\bar{x}} \cdot (1-p)^{n-n\bar{x}}}$$

Problem 2: Verify the joint pdf of the normal sample on page 9 of Slides02 and show that it depends on the sample only through the first two sample moments.

Slides02, page 7

Normal population $X \sim \mathcal{N}(\mu, \sigma^2)$ The cdf and pdf are

$$F_1(x, \boldsymbol{\theta}) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y-u)^2}{2\sigma^2}}\, dy, \qquad f_1(x, \boldsymbol{\theta}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-u)^2}{2\sigma^2}},$$

respectively, and the two-dimensional parameter $\boldsymbol{\theta} = (\mu, \sigma^2)'$. □

Slides02, page 9

If $X \sim \mathcal{N}(\mu, \sigma^2)$, then $(X_1, \ldots, X_n)$ has joint pdf

$$\prod_{i=1}^{n} f_1(x_i, \boldsymbol{\theta}) = \frac{1}{(2\pi\sigma^2)^{n/2}} \exp\left\{ \frac{-1}{2\sigma^2} \left[\sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2 \right] \right\},$$

where $\boldsymbol{\theta} = (\mu, \sigma^2)'$ and $x_i \in (-\infty, \infty)$, $i = 1, \ldots, n$.

$$f_1(x, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{\frac{-(x-u)^2}{2\sigma^2}}$$

$$\theta = (\mu, \sigma^2)$$

$$\begin{array}{cc} \downarrow & \downarrow \\ \text{mean} & \text{variance} \\ \downarrow & \downarrow \\ 1^{\text{st}} \text{ sample moment} & 2^{\text{nd}} \text{ sample moment} \end{array}$$

$$\prod_{i=1}^{n} f_i(x_i, \theta) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_1-\mu)^2}{2\sigma^2}} \cdot \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x_2-\mu)^2}{2\sigma^2}} \cdot \ldots$$

take it out of the product b/c its constant

$$= \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \prod_{i=1}^{n} e^{-\frac{(x_i-\mu)^2}{2\sigma^2}}$$

knowing

$$\prod_{b=1}^{z} a^b = a^{\sum_{b=1}^{z} b},$$

lets say $a = e$, $b = -\frac{(x_i-\mu)^2}{2\sigma^2}$

$$\Rightarrow \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{\sum_{i=1}^{n}\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)}$$

we can simplify further by taking out the constants in the summation

$$\Rightarrow \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot e^{\left(-\frac{1}{2\sigma^2}\right) \cdot \underline{\sum_{i=1}^{n} (x_i-\mu)^2}}$$

$$\sum_{i=1}^{n} (x_i-\mu)^2 = \sum_{i=1}^{n} (x_i^2 - 2\mu x_i + \mu^2) = \sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2$$

$$\sum_{i=1}^{n} x_i^2 \qquad -2\mu \sum_{i=1}^{n} x_i \qquad \sum_{i=1}^{n} \mu^2$$

$\hookrightarrow n\mu^2$

sub in the summation

$$\Rightarrow \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot \exp\left(-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} x_i^2 - 2\mu \sum_{i=1}^{n} x_i + n\mu^2\right]\right)$$

$\exp(j) = e^j$

we know

$i=1 \qquad i=1$

$\sum_{i=1}^{n} x_i = n\bar{x}$ — $\bar{x}$ = sample mean = 1st sample moment

$\sum_{i=1}^{n} x_i^2 = n\sigma^2$ — σ^2 = variance = 2nd sample moment

$$\Rightarrow \left(\frac{1}{\sqrt{2\pi\sigma^2}}\right)^n \cdot \exp\left(-\frac{1}{2\sigma^2}\left[n\sigma^2 - 2\mu n\bar{x} + n\mu^2\right]\right)$$

Problem 3: Suppose X ~ N (μ, σ2). Show that (X−μ)/σ ~ N (0, 1), the standardization on page 20 of Slides02.

Slides02, page 19

✎ A r.v. X is of normal distribution with parameter (μ, σ^2), denoted by $X \sim \mathcal{N}(\mu, \sigma^2)$, if it has pdf

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}},$$

where $x \in \mathbb{R}$, $\mu \in \mathbb{R}$ and $\sigma^2 > 0$.

Slides02, page 20

- The mean and variance are

$$\mathrm{E}[X] = \mu \quad \text{and} \quad \mathrm{Var}[X] = \sigma^2, \quad \text{respectively.}$$

- For the standard normal distribution $\mathcal{N}(0, 1)$, the pdf and cdf are respectively denoted by

$$\phi(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \qquad \Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \mathrm{dy}.$$

- For $X \sim \mathcal{N}(\mu, \sigma^2)$, the standardization

$$\frac{X-\mu}{\sigma} \sim \mathcal{N}(0, 1).$$

$\frac{(x-\mu)}{\sigma} \sim N(0,1)$ $\qquad \frac{dx}{dz} = \sigma$

$\Rightarrow z = \frac{x-\mu}{\sigma} = x = \sigma z + \mu$

$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$

sub in x $\Rightarrow \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\sigma z + \cancel{\mu - \mu})^2}{2\sigma^2}} \cdot \sigma$

$= \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{z^2}{2}} \cdot \sigma$

$= \frac{1}{\sqrt{2\pi\cancel{\sigma^2}}} \cdot \cancel{\sigma} \cdot e^{-\frac{z^2}{2}}$

$= \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$

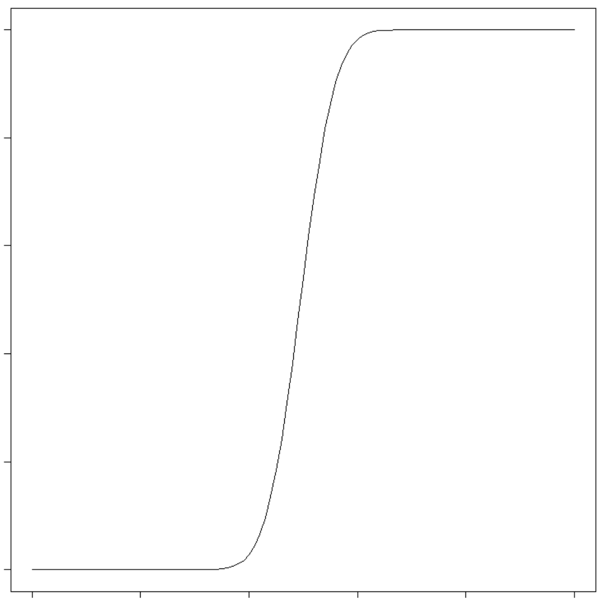
Problem 4

1. The original population is the entire block of goods.
2. $X = \{X_1, \dots, X_n\}$ where $X_i = 1\ if\ defective,\ 0\ if\ qualified,\ where\ i = 1, \dots, n.$
3. $X \sim B(1, \frac{d}{d+m})$, which means its cdf is:

   $F(x, \frac{d}{d+m}) = 0, \quad if\ x < 0,$

   $= 1 - \frac{d}{d+m}, \quad if\ 0 \le x \le 1,$

   $= 1, \quad if\ x \ge 1$



Problem 5

1. $X_i = 1\ if\ defective,\ 0\ if\ qualified,\ where\ i = 1, \dots, n.$
2. No. When chosen with replacement, the sample is no longer a SRS.
3. The probability does not change per good, so $p = \frac{d}{d+m}$.
4. $N_1 \sim B(d + m, \frac{d}{d+m})$
5. $E[N_1] = np = (d + m)(\frac{d}{d+m}) = d$

   $Var[N_1] = np(1 - p) = d(1 - \frac{d}{d+m}) = \frac{dm}{d+m}$

Problem 6

1. Yes. Assuming they were randomly selected and each good is mutually independent, drawing with replacement will yield a SRS.
2. On the second trial, $p = \frac{d}{d+m-1}$ or $\frac{d-1}{d+m-1}$, depending on if the first trial yielded a qualified or defective good, respectively.
3. $N_2 \sim Hypergeometric(d + m, d, 1)$

4. $E[N_2] = (d + m)\frac{d}{d+m} = d$

   $Var[N_2] = n\frac{K}{N}\frac{N-K}{N}\frac{N-n}{N-1} = \frac{d}{d+m}\frac{m}{d+m} = \frac{dm}{(d+m)^2}$

Problem 7

1. $N_1 = \{X_1, \dots, X_n\}$, where some Xs may be observations of the same goods.

   $N_2 = \{X_1, \dots, X_n\}$, where all of the Xs will be observations of different goods.
2. $E[N_2] = d = E[N_1] = d$. Both formulas are equal to np. It makes sense because the data doesn't change, no matter how we sample it.
3. $Var[N_2] = \frac{dm}{(d+m)^2} \neq Var[N_1] = \frac{dm}{d+m}$. The way we sample the data has changed, so it makes sense that the variance changes.