

HW1

1.5-4

(a). Show that T_1 and T_2 are equivalent statistics if, and only if, we can write $T_2 = H(T_1)$ for some 1-1 transformation H of the range of T_1 into the range of T_2 . Which of the following statistics are equivalent? (Prove or disprove.)

If $T_2 = H(T_1)$ for some 1-1 transformation H of the range of T_1 into the range of T_2 , then

when $T_1(x) = T_1(y)$, $T_2(x) = H(T_1(x)) = H(T_1(y)) = T_2(y)$;

when $T_2(x) = T_2(y)$, $H(T_1(x)) = T_2(x) = T_2(y) = H(T_1(y))$; then T_1 and T_2 are equivalent.

If T_1 and T_2 are equivalent, then $\exists H$ make $T_2 = H(T_1)$ is a 1-1 transformation of the range of T_1 into the range of T_2 .

Therefore, T_1 and T_2 are equivalent statistics $\iff T_2 = H(T_1)$. ■

(b). $\prod_{i=1}^n x_i$ and $\sum_{i=1}^n \log x_i$, $x_i > 0$

$T_2(x) = \sum_{i=1}^n \ln x_i = \ln(\prod_{i=1}^n x_i) = \ln(T_1)$, $x_i > 0$. $H(x) = \ln x$ is a 1-1 transformation of $T_1 \in (0, \infty)$ into $T_2 \in (-\infty, \infty)$.

Thus, T_1 and T_2 are equivalent. ■

(c). $\sum_{i=1}^n x_i$ and $\sum_{i=1}^n \log x_i$, $x_i > 0$

$T_2(x) = \sum_{i=1}^n \ln x_i = T_1(\ln(x)) \neq T_1(x)$, $x_i > 0$.

There is not a H that can do a 1-1 transformation of the range of T_1 into the range of T_2 .

Thus, T_1 and T_2 are not equivalent. ■

(d). $(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^2)$ and $(\sum_{i=1}^n x_i, \sum_{i=1}^n (x_i - \bar{x})^2)$

Let $T_1 = (T_{11} = \sum_{i=1}^n x_i, T_{12} = \sum_{i=1}^n x_i^2)$, then

$$T_{21} = \sum_{i=1}^n x_i = T_{11}$$

$$T_{22} = \sum_{i=1}^n (x_i - \bar{x})^2 = \sum_{i=1}^n x_i^2 - 2\bar{x} \sum_{i=1}^n x_i + n(\bar{x})^2 = \sum_{i=1}^n x_i^2 - \frac{2}{n} (\sum_{i=1}^n x_i)^2 + \frac{1}{n} (\sum_{i=1}^n x_i)^2 = T_{12} - \frac{1}{n} T_{11}^2$$

H is a 1-1 transformation of the range of T_1 into the range of T_2 . Thus, T_1 and T_2 are equivalent. ■

(e). $(\sum_{i=1}^n x_i, \sum_{i=1}^n x_i^3)$ and $(\sum_{i=1}^n x_i, \sum_{i=1}^n (x_i - \bar{x})^3)$

Let $T_1 = (T_{11} = \sum_{i=1}^n x_i, T_{12} = \sum_{i=1}^n x_i^3)$, then

$$T_{21} = \sum_{i=1}^n x_i = T_{11}$$

$$\begin{aligned} T_{22} &= \sum_{i=1}^n (x_i - \bar{x})^3 = \sum_{i=1}^n x_i^3 - 3\bar{x} \sum_{i=1}^n x_i^2 + 3\bar{x}^2 \sum_{i=1}^n x_i - n(\bar{x})^3 = \\ &= \sum_{i=1}^n x_i^3 - \frac{3}{n} \sum_{i=1}^n x_i \sum_{i=1}^n x_i^2 + \frac{3}{n^2} (\sum_{i=1}^n x_i)^3 - \frac{1}{n^2} (\sum_{i=1}^n x_i)^3 = T_{12} - \frac{3}{n} T_{11} \sum_{i=1}^n x_i^2 + \frac{2}{n^2} T_{11}^3 \end{aligned}$$

There is not statistics in T_1 can represent $\sum_{i=1}^n x_i^2$. There is not a H that can do a 1-1 transformation of the range of T_1 into the range of T_2 . Thus, T_1 and T_2 are not equivalent. ■

1.5-6 Let X take on the specified values v_1, \dots, v_k with probabilities $\theta_1, \dots, \theta_k$, respectively.

Suppose that X_1, \dots, X_n are independently and identically distributed as X . Suppose that $\theta = (\theta_1, \dots, \theta_k)$ is unknown and may range over the set $\Theta = \{(\theta_1, \dots, \theta_k) : \theta_i \geq 0, 1 \leq i \leq k, \sum_{i=1}^k \theta_i = 1\}$. Let N_j be the number of X_i which equal v_j .

(a). What is the distribution of (N_1, \dots, N_k) ?

$(N_1, \dots, N_k) \sim \text{Multinomial Distribution}$

$$f_{\vec{\theta}}(\vec{n}) = n! \prod_{i=1}^k \frac{\theta_i^{n_i}}{n_i!} \mathbf{1}_{\{\sum N_i = n\}}, \text{ where } n_i = \text{the number of times we get outcome } i = 1, \dots, k \quad \blacksquare$$

(b). Show that $\mathbf{N} = (N_1, \dots, N_{k-1})$ is sufficient for θ .

$$f_{\vec{\theta}}(\vec{N}) = n! \prod_{i=1}^k (N_i!)^{-1} \exp[\sum_{i=1}^k N_i \ln \theta_i] \mathbf{1}_{\{\sum N_i = n\}} = h(\vec{N}) \exp[\sum_{i=1}^k \eta_i(\vec{\theta}) T_i(\vec{N}) - B(\vec{\theta})], \text{ where } \chi = \{\vec{N} \in \{0, \dots, n\}^k \mid \sum N_i = n\}$$

$$h(\vec{N}) = n! \prod_{i=1}^k (N_i!)^{-1} \mathbf{1}_{\{\sum N_i = n\}}, \quad B(\vec{\theta}) = 0$$

$$\eta_i(\vec{\theta}) = (\ln \theta_1, \dots, \ln \theta_k),$$

$T(\vec{N}) = (N_1, \dots, N_k)$ is a n.s.s of the family.

$T(\vec{N}) = (N_1, \dots, N_{k-1}, n - \sum_{i=1}^{k-1} N_i)$ is equivalent with (N_1, \dots, N_{k-1}) . Therefore \mathbf{N} is sufficient for θ . ■

1.5-7 Let X_1, \dots, X_n be a sample from a population with density $p(x, \theta)$ given by

$$p(x, \theta) = \begin{cases} \frac{1}{\sigma} \exp\{-\frac{x-\mu}{\sigma}\} & \text{if } x \geq \mu \\ 0 & \text{o.w.} \end{cases} \quad \text{Here } \theta = (\mu, \sigma) \text{ with } -\infty < \mu < \infty, \sigma > 0.$$

(a) Show that $\min(X_1, \dots, X_n)$ is sufficient for μ when σ is fixed.

When σ is fixed, $p(x_{1:n}, \mu) = \sigma^{-n} \exp[-\frac{\sum_{i=1}^n x_i}{\sigma}] \exp[\frac{n\mu}{\sigma}] \prod_{i=1}^n \mathbf{1}_{\{x_i \geq \mu\}}$, where

$$h(x) = \sigma^{-n} \exp[-\frac{\sum_{i=1}^n x_i}{\sigma}], \quad g(T(x), \mu) = \exp[\frac{n\mu}{\sigma}] \prod_{i=1}^n \mathbf{1}_{\{x_i \geq \mu\}}$$

$\mathbf{1}_{\{x_{(1)} \geq \mu\}}$ contains all the information about μ , then

$T(x) = \min(X_1, \dots, X_n)$ is sufficient for μ when σ is fixed. ■

- Another method is that $p(x_{1:n}|t)$ is free of μ

$$X \sim \text{Expo}(\mu, 1/\sigma), \quad F_{\mu, \sigma}(x) = 1 - e^{-(x-\mu)/\sigma},$$

$$\min(X_1, \dots, X_n) = X_{(1)} = n \frac{1}{\sigma} e^{-(x-\mu)/\sigma} [1 - (1 - e^{-(x-\mu)/\sigma})]^{n-1} = \frac{n}{\sigma} e^{-n(x-\mu)/\sigma}$$

$$p(x_{1:n}|t) = \frac{1}{n\sigma^{n-1}} e^{\frac{1}{\sigma}(\sum x_i - nx)} \text{ is free of } \mu$$

(b) Find a one-dimensional sufficient statistic for σ when μ is fixed.

When μ is fixed, $p(x_{1:n}, \sigma) = \sigma^{-n} \exp[-\frac{\sum_{i=1}^n x}{\sigma} + \frac{n\mu}{\sigma}] \prod_{i=1}^n \mathbf{1}_{\{x_i \geq \mu\}}$, where

$$h(x) = \prod_{i=1}^n \mathbf{1}_{\{x \geq \mu\}},$$

$$g(T(x), \sigma) = \sigma^{-n} \exp[-\frac{\sum_{i=1}^n x}{\sigma} + \frac{n\mu}{\sigma}], \text{ then}$$

$T(x) = \sum_{i=1}^n x_i$ is sufficient for σ when μ is fixed. ■

- Another method is that $p(x_{1:n}|t)$ is free of σ

$$X \sim \text{Exp}(\mu, 1/\sigma), F_{\mu, \sigma}(x) = 1 - e^{-(x-\mu)/\sigma},$$

$$Y = X - \mu \sim \text{Exp}(1/\sigma), T = \sum Y_i \sim \text{Gamma}(n, \sigma)$$

$$p(x_{1:n}|t) = \Gamma(n)t^{1-n} \text{ is free of } \sigma$$

(c) Exhibit a two-dimensional sufficient statistic for θ .

$$p(x_{1:n}, \mu, \sigma) = \sigma^{-n} \exp[-\frac{\sum_{i=1}^n x}{\sigma} + \frac{n\mu}{\sigma}] \prod_{i=1}^n \mathbf{1}_{\{x_i \geq \mu\}}, \text{ where}$$

$$h(x) = 1,$$

$$g(T(x), \mu, \sigma) = \sigma^{-n} \exp[-\frac{\sum_{i=1}^n x}{\sigma} + \frac{n\mu}{\sigma}] \prod_{i=1}^n \mathbf{1}_{\{x_{(1)} \geq \mu\}}, \text{ then}$$

$$T(x) = (x_{(1)}, \sum_{i=1}^n x_i) \text{ is a two-dimensional sufficient statistic for } \theta. \quad \blacksquare$$

1.5-9 Let X_1, \dots, X_n be a sample from a population with density

$$f_{\theta}(x) = \begin{cases} a(\theta)h(x) & \text{if } \theta_1 \leq x \leq \theta_2 \\ 0 & \text{o.w.} \end{cases} \text{ where } h(x) \geq 0, \theta = (\theta_1, \theta_2) \text{ with } -\infty < \theta_1 \leq \theta_2 < \infty, \text{ and } a(\theta) =$$

$[\int_{\theta_1}^{\theta_2} h(x)dx]^{-1}$ is assumed to exist. Find a two-dimensional sufficient statistic for this problem and apply your result to the $U[\theta_1, \theta_2]$ family of distributions.

$$\text{Let } H'(x) = h(x), a(\theta) = [\int_{\theta_1}^{\theta_2} h(x)dx]^{-1} = [H(\theta_2) - H(\theta_1)]^{-1}$$

$$f_{\theta_1, \theta_2}(x_{1:n}) = \prod_{i=1}^n [a(\theta)h(x)\mathbf{1}_{\{x \in [\theta_1, \theta_2]\}}] = \prod_{i=1}^n [\mathbf{1}_{\{x \in [\theta_1, \theta_2]\}}][H(\theta_2) - H(\theta_1)]^{-n} \prod_{i=1}^n h(x), \text{ where}$$

$$g(T(x), \theta_1, \theta_2) = \prod_{i=1}^n [\mathbf{1}_{\{x \in [\theta_1, \theta_2]\}}][H(\theta_2) - H(\theta_1)]^{-n},$$

$$h'(x) = \prod_{i=1}^n h(x)$$

$\mathbf{1}_{\{x_{(n)} \leq \theta_2\}} \mathbf{1}_{\{x_{(1)} \geq \theta_1\}}$ contains all the information about θ , then

$$T(x) = (x_{(1)}, x_{(n)}) \text{ is a two-dimensional sufficient statistic for } \theta. \quad \blacksquare$$

For $U[\theta_1, \theta_2]$, let $h(x) = 1$, $a(\theta) = (\theta_2 - \theta_1)^{-1}$

$$f_{\theta_1, \theta_2}(x_{1:n}) = \prod_{i=1}^n [a(\theta)h(x)\mathbf{1}_{\{x \in [\theta_1, \theta_2]\}}] = \prod_{i=1}^n [\mathbf{1}_{\{x \in [\theta_1, \theta_2]\}}][\theta_2 - \theta_1]^{-n} \prod_{i=1}^n 1, \text{ where}$$

$$g(T(x), \theta_1, \theta_2) = \prod_{i=1}^n [\mathbf{1}_{\{x_{(n)} \leq \theta_2\}} \mathbf{1}_{\{x_{(1)} \geq \theta_1\}}][\theta_2 - \theta_1]^{-n},$$

$$h'(x) = 1$$

$$T(x) = (x_{(1)}, x_{(n)}) \text{ is a two-dimensional sufficient statistic for } \theta \text{ in the } U[\theta_1, \theta_2] \text{ family.} \quad \blacksquare$$

HW2

1.6-1 Prove the assertions of Table 1.6.1

		$\eta(\theta)$	$T(x)$
$N(\mu, \sigma^2)$	σ^2 fixed	μ/σ^2	x
	μ fixed	$-1/2\sigma^2$	$(x - \mu)^2$
$\Gamma(p, \lambda)$	p fixed	$-\lambda$	x
	λ fixed	$(p - 1)$	$\log x$
$\beta(r, s)$	r fixed	$(s - 1)$	$\log(1 - x)$
	s fixed	$(r - 1)$	$\log x$

For Normal distribution,

$$f_{\mu}(x) = \exp\left[\underbrace{\frac{\mu}{\sigma^2}}_{\eta(\mu)} \underbrace{x}_{T(x)} - \underbrace{\left(\frac{\mu^2}{2\sigma^2} + \ln(\sqrt{2\pi}\sigma)\right)}_{B(\mu)} \underbrace{\exp\left[-\frac{x^2}{2\sigma^2}\right]}_{h(x)} \mathbf{1}_{\{x \in \mathbb{R}\}}\right] \quad (\text{When } \sigma^2 \text{ fixed})$$

$$f_{\sigma^2}(x) = \exp\left[\underbrace{-\frac{1}{2\sigma^2}}_{\eta(\sigma^2)} \underbrace{(x - \mu)^2}_{T(x)} - \underbrace{\ln(\sqrt{2\pi}\sigma)}_{B(\sigma^2)} \underbrace{\mathbf{1}_{\{x \in \mathbb{R}\}}}_{h(x)}\right] \quad (\text{When } \mu \text{ fixed})$$

For Gamma distribution,

$$f_{\lambda}(x) = \exp\left[\underbrace{-\lambda}_{\eta(\lambda)} \underbrace{x}_{T(x)} - \underbrace{\ln\left(\frac{\lambda^p}{\Gamma p}\right)}_{B(\lambda)} \underbrace{x^{p-1} \mathbf{1}_{\{x \in (0, \infty)\}}}_{h(x)}\right] \quad (\text{When } p \text{ fixed})$$

$$f_p(x) = \exp\left[\underbrace{(p - 1)}_{\eta(p)} \underbrace{\ln(x)}_{T(x)} - \underbrace{\ln\left(\frac{\lambda^p}{\Gamma p}\right)}_{B(p)} \underbrace{\exp[-\lambda x] \mathbf{1}_{\{x \in (0, \infty)\}}}_{h(x)}\right] \quad (\text{When } \lambda \text{ fixed})$$

For Beta distribution,

$$f_s(x) = \exp\left[\underbrace{(s - 1)}_{\eta(s)} \underbrace{\ln(1 - x)}_{T(x)} - \underbrace{\ln(B(r, s))}_{B(s)} \underbrace{x^{r-1} \mathbf{1}_{\{x \in (0, 1)\}}}_{h(x)}\right] \quad (\text{When } r \text{ fixed})$$

$$f_r(x) = \exp\left[\underbrace{(r - 1)}_{\eta(r)} \underbrace{\ln(x)}_{T(x)} - \underbrace{\ln(B(r, s))}_{B(r)} \underbrace{(1 - x)^{s-1} \mathbf{1}_{\{x \in (0, 1)\}}}_{h(x)}\right] \quad (\text{When } s \text{ fixed})$$

1.6-3 Let X be the number of failures before the first success in a sequence of Bernoulli trials

with probability of success θ . Then $P_\theta[X = k] = (1 - \theta)^k \theta, k = 0, 1, 2, \dots$. This is called the geometric distribution ($G(\theta)$).

(a) Show that the family of geometric distributions is a one-parameter exponential family with $T(x) = x$.

For Geometric distribution,

$$P_\theta(X = k) = \exp \underbrace{[\ln(1 - \theta)]}_{\eta(\theta)} \underbrace{k}_{T(k)} - \underbrace{\ln(\theta)}_{B(\theta)} \underbrace{\mathbf{1}_{\{k \in (0, 1, 2, \dots)\}}}_{h(k)}$$

Thus, geometric distributions is a one-parameter exponential family with $T(x) = x$

(b) Deduce from Theorem 1.6.1 that if X_1, \dots, X_n is a sample from $G(\theta)$, then the distributions of $\sum_{i=1}^n X_i$ form a one-parameter exponential family.

$$P_\theta(X_{1:n}) = \prod_{i=1}^n P_\theta[X = x] = \exp \underbrace{[\ln(1 - \theta)]}_{\eta(\theta)} \underbrace{\sum_{i=1}^n x_i}_{T(x)} - \underbrace{n \ln(\theta)}_{B(\theta)} \underbrace{\prod_{i=1}^n \mathbf{1}_{\{x \in (0, 1, 2, \dots)\}}}_{h(x)}$$

$\sum_{i=1}^n X_i$ is a sufficient statistic for θ for a one-parameter exponential family. By theorem 1.6.1, the family of the distribution of $\sum_{i=1}^n X_i$ is a one-parameter exponential family, whose p.m.f may be written as $h^*(t) \exp[\eta(\theta)t - B(\theta)]$ for a suitable h^* .

(c) Show that $\sum_{i=1}^n X_i$ in part (b) has a negative binomial distribution with parameters (n, θ) defined by $P_\theta[\sum_{i=1}^n X_i = k] = \binom{n+k-1}{k} (1 - \theta)^k \theta^n, k = 0, 1, 2, \dots$ (The negative binomial distribution is that of the number of failures before the n th success in a sequence of Bernoulli trials with probability of success θ .) Hint: By Theorem 1.6.1, $P_\theta[\sum_{i=1}^n X_i = k] = c_k (1 - \theta)^k \theta^n, 0 < \theta < 1$. If $\sum_{k=0}^{\infty} c_k \omega^k = \frac{1}{(1 - \omega)^n}, 0 < \omega < 1$, then $c_k = \frac{1}{k!} \frac{d^k}{d\omega^k} (1 - \omega)^{-n} \Big|_{\omega=0}$

To find p.m.f of this distribution, let $\sum_{k=1}^n c_k (1 - \theta)^k \theta^n = 1, 0 < \theta < 1$

let $\omega = 1 - \theta, \sum_{k=1}^n c_k \omega^k = \theta^{-n}, 0 < \omega < 1$, then $c_k = \frac{1}{k!} \frac{d^k}{d\omega^k} (1 - \omega)^{-n} \Big|_{\omega=0}$

$$\begin{aligned} \frac{d'}{d\omega'} (1 - \omega)^{-n} &= (-n)(-1)(1 - \omega)^{-n-1} = n(1 - \omega)^{-n-1} \\ \frac{d^2}{d\omega^2} (1 - \omega)^{-n} &= (-n-1)(-1)n(1 - \omega)^{-n-2} = (n+1)n(1 - \omega)^{-n-2} \\ &\dots \end{aligned}$$

$$\frac{d^k}{d\omega^k} (1 - \omega)^{-n} = (-n-k+1)(-1) \dots (n+1)n(1 - \omega)^{-n-k} = \left[\prod_{i=1}^k (n+i-1) \right] (1 - \omega)^{-n-k}$$

$$\frac{d^k}{d\omega^k} (1 - \omega)^{-n} \Big|_{\omega=0} = \prod_{i=1}^k (n+i-1) = \prod_{i=0}^{k-1} (n+i)$$

$$c_k = \frac{1}{k!} \frac{d^k}{d\omega^k} (1 - \omega)^{-n} \Big|_{\omega=0} = \frac{1}{k!} \prod_{i=0}^{k-1} (n+i) = \frac{(n+k-1)!}{k!(n-1)!} = \binom{n+k-1}{k}$$

Therefore, $P_\theta[\sum_{i=1}^n X_i = k] = \binom{n+k-1}{k} (1 - \theta)^k \theta^n, k = 0, 1, 2, \dots$

1.6-5 Show that the following families of distributions are two-parameter exponential families and identify the functions η, B, T , and h .

(a) The beta family.

$$f_{r,s}(x) = \exp\left[\underbrace{(r-1)\ln(x) + (s-1)\ln(1-x)}_{\eta(r,s)T(x)} - \underbrace{\ln(B(r,s))}_{B(r,s)}\right] \underbrace{\mathbf{1}_{\{x \in (0,1)\}}}_{h(x)}$$

The beta family is a two-parameter exponential family with $\eta(r,s) = (r-1, s-1)^T$; $T(x) = (\ln(x), \ln(1-x))$; $B(r,s) = \frac{\Gamma(r+s)}{\Gamma(r)\Gamma(s)}$; $h(x) = \mathbf{1}_{\{x \in (0,1)\}}$

(b) The gamma family.

$$f_{p,\lambda}(x) = \exp\left[\underbrace{-\lambda x + (p-1)\ln x}_{\eta(p,\lambda)T(x)} - \underbrace{\ln\left(\frac{\lambda^p}{\Gamma p}\right)}_{B(p,\lambda)}\right] \underbrace{\mathbf{1}_{\{x \in (0,\infty)\}}}_{h(x)}$$

The gamma family is a two-parameter exponential family with $\eta(p,\lambda) = (-\lambda, (p-1))^T$; $T(x) = (x, \ln(x))$; $B(p,\lambda) = -\ln\left(\frac{\lambda^p}{\Gamma p}\right)$; $h(x) = \mathbf{1}_{\{x \in (0,\infty)\}}$

1.6-7 Let $\mathbf{X} = ((X_1, Y_1), \dots, (X_n, Y_n))$ be a sample from a bivariate normal population.

Show that the distributions of \mathbf{X} form a five-parameter exponential family and identify η, B, T , and h .

$$\begin{aligned} f(\vec{X}, \vec{Y}) &= \exp\left[-\frac{1}{2(1-\rho^2)}\left[\sum_{i=1}^n \left(\frac{x_i - \mu_X}{\sigma_X}\right)^2 - 2\rho \sum_{i=1}^n \left(\frac{x_i - \mu_X}{\sigma_X}\right)\left(\frac{y_i - \mu_Y}{\sigma_Y}\right) + \sum_{i=1}^n \left(\frac{y_i - \mu_Y}{\sigma_Y}\right)^2\right] - n \ln(2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})\right] \mathbf{1}_{\{x,y \in \mathbb{R}^n\}} \\ &= \exp\left[\underbrace{-\frac{\sum x^2}{2(1-\rho^2)\sigma_X^2} + \frac{\sum x}{(1-\rho^2)}\left(\frac{\mu_X}{\sigma_X^2} - \frac{\mu_Y\rho}{\sigma_X\sigma_Y}\right) + \frac{\rho \sum xy}{(1-\rho^2)\sigma_X\sigma_Y} + \frac{\sum y}{(1-\rho^2)}\left(\frac{\mu_Y}{\sigma_Y^2} - \frac{\mu_X\rho}{\sigma_X\sigma_Y}\right) - \frac{\sum y^2}{2(1-\rho^2)\sigma_Y^2}}_{\eta(\rho, \mu_X, \mu_Y, \sigma_X, \sigma_Y)T(x,y)}\right] \\ &\quad \cdot \exp\left[\underbrace{-n\left(\frac{1}{2(1-\rho^2)}\left(\frac{\mu_X^2}{\sigma_X^2} - \frac{2\rho\mu_X\mu_Y}{\sigma_X\sigma_Y} + \frac{\mu_Y^2}{\sigma_Y^2}\right) + \ln(2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})\right)}_{nB(\rho, \mu_X, \mu_Y, \sigma_X, \sigma_Y)}\right] \underbrace{\mathbf{1}_{\{x,y \in \mathbb{R}^n\}}}_{h(x)} \end{aligned}$$

where

$$\eta(\rho, \mu_X, \mu_Y, \sigma_X, \sigma_Y) = \left\{ -\frac{1}{2(1-\rho^2)\sigma_X^2}, \frac{1}{(1-\rho^2)}\left(\frac{\mu_X}{\sigma_X^2} - \frac{\mu_Y\rho}{\sigma_X\sigma_Y}\right), \frac{\rho}{(1-\rho^2)\sigma_X\sigma_Y}, \frac{1}{(1-\rho^2)}\left(\frac{\mu_Y}{\sigma_Y^2} - \frac{\mu_X\rho}{\sigma_X\sigma_Y}\right), -\frac{1}{2(1-\rho^2)\sigma_Y^2} \right\}^T$$

$$T(x, y) = (\sum x^2, \sum x, \sum xy, \sum y, \sum y^2); h(x) = \mathbf{1}_{\{x,y \in \mathbb{R}^n\}}$$

$$nB(\rho, \mu_X, \mu_Y, \sigma_X, \sigma_Y) = n\left(\frac{1}{2(1-\rho^2)}\left(\frac{\mu_X^2}{\sigma_X^2} - \frac{2\rho\mu_X\mu_Y}{\sigma_X\sigma_Y} + \frac{\mu_Y^2}{\sigma_Y^2}\right) + \ln(2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2})\right)$$

$$\rho \in (0, 1), \mu_X \in \mathbb{R}, \mu_Y \in \mathbb{R}, \sigma_X \in \mathbb{R}^+, \sigma_Y \in \mathbb{R}^+$$

$$x \in \mathcal{X} \subset \mathbb{R}^n, y \in \mathcal{Y} \subset \mathbb{R}^n$$

HW3

2.1-1 Consider a population made up of three different types of individuals occurring in the Hardy-Weinberg proportions θ^2 , $2\theta(1-\theta)$ and $(1-\theta)^2$, respectively, where $0 < \theta < 1$.

$$\begin{array}{c|ccc|c} j & 1 & 2 & 3 & \\ N_j & N_1 & N_2 & N_3 & \sum N_j = n \\ p_j & p_1 = \theta^2 & p_2 = 2\theta(1-\theta) & p_3 = (1-\theta)^2 & \\ \hat{p}_j & N_1/n & N_2/n & N_3/n & \sum \hat{p}_j = 1 \\ x_j & -1 & 0 & 1 & \end{array}$$

(a) Show that $T_3 = N_1/n + N_2/2n$ is a frequency substitution estimate of θ .

$$E[T_3] = E[N_1/n + N_2/2n] = E[\hat{p}_1] + \frac{1}{2}E[\hat{p}_2] = \theta^2 + \frac{1}{2} \cdot 2\theta(1-\theta) = \theta$$

$\therefore T_3$ is a frequency substitution estimate of θ .

(b) Using the estimate of (a), what is a frequency substitution estimate of the odds ratio $\frac{\theta}{1-\theta}$?

$$\text{For } \hat{p}_3 = 1 - \frac{N_1}{n} - \frac{N_2}{n}$$

$$\text{Let } g\left(\frac{N_1}{n}, \frac{N_2}{n}\right) = \left(1 - \frac{N_1}{n} - \frac{N_2}{n}\right)^{-\frac{1}{2}} - 1 = (\hat{p}_3)^{-\frac{1}{2}} - 1 \quad \text{convert to 1-parametric function.}$$

$$E[g(\frac{N_1}{n}, \frac{N_2}{n})] = E[\frac{1}{\sqrt{\hat{p}_3}} - 1] = \frac{1}{\sqrt{E[\hat{p}_3]}} - 1 = \frac{1}{\sqrt{(1-\theta)^2}} - 1 \stackrel{0 < \theta < 1}{=} \frac{\theta}{1-\theta}$$

$$T(X_1, X_2) = g(\frac{N_1}{n}, \frac{N_2}{n}) \text{ is a frequency substitution estimate of the odds ratio } \frac{\theta}{1-\theta}$$

(c) Suppose X takes the values $-1, 0, 1$ with respective probabilities p_1, p_2, p_3 given by the Hardy-Weinberg proportions. By considering the first moment of X , show that T_3 is a method of moment estimate of θ .

$$\mu_1(\theta) = E_\theta[X^1] = \sum_{j=1}^3 p_j x_j = p_1 x_1 + p_2 x_2 + p_3 x_3 = \theta^2 \cdot (-1) + 2\theta(1-\theta) \cdot 0 + (1-\theta)^2 \cdot 1 = 1 - 2\theta$$

$$\hat{\mu}_1 = \frac{1}{n} \sum_{i=1}^n x_i^1 = \frac{1}{n} (N_1 x_1 + N_2 x_2 + N_3 x_3) = \frac{1}{n} [N_1 \cdot (-1) + N_2 \cdot 0 + (n - N_1 - N_2) \cdot 1] = 1 - 2\left(\frac{N_1}{n} + \frac{N_2}{2n}\right) = 1 - 2T_3$$

Since $E[\hat{\mu}_1] = 1 - 2E[T_3] = 1 - 2\theta = \mu_1$, T_3 is a method of moment estimate of θ .

2.1-9 Suppose $\mathbf{X} = (X_1, \dots, X_n)$ where the X_i are independent $N(0, \sigma^2)$.

(a) Find an estimate of σ^2 based on the second moment.

$$\mu_2(\sigma^2) = E_{\sigma^2}[X^2] = V[X] + E[X] = \sigma^2$$

$$E[\hat{\mu}_2] = E\left[\frac{1}{n} \sum_{i=1}^n x_i^2\right] = \mu_2(\sigma^2) = \sigma^2$$

Therefore, $\frac{1}{n} \sum_{i=1}^n x_i^2$ is an estimate of σ^2 based on the second moment.

(b) Construct an estimate of σ using the estimate of part (a) and the equation $\sigma = \sqrt{\sigma^2}$.

Let $T_2 = \sqrt{\hat{\mu}_2} = \sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$ is a 1-1 transformation in $(0, \infty)$. Therefore T_2 and $\hat{\mu}_2$ are equivalent statistics.

$$E[T_2] = E[\sqrt{\hat{\mu}_2}] = \sqrt{E[\hat{\mu}_2]} = \sqrt{E\left[\frac{1}{n} \sum_{i=1}^n x_i^2\right]} = \sqrt{\sigma^2} = \sigma, \quad x \in (0, \infty)$$

Therefore, $\sqrt{\frac{1}{n} \sum_{i=1}^n x_i^2}$ is an estimate of σ based on the second moment.

(c) Use the empirical substitution principle to construct an estimate of σ using the relation $E(|X_1|) = \sigma\sqrt{2/\pi}$.

If $X \sim N(0, \sigma^2)$, then $|X| \sim \text{half-normal}(\sigma\sqrt{\frac{2}{\pi}}, \sigma^2(1 - \frac{2}{\pi}))$. We have $\mu_1(\sigma) = E_\sigma[|X_1|] = \sigma\sqrt{\frac{2}{\pi}}$

Let $T_1 = \sqrt{\frac{\pi}{2}}\hat{\mu}_1 = \sqrt{\frac{\pi}{2}}\frac{1}{n} \sum_{i=1}^n |x_i|$ is a 1-1 transformation in \mathbb{R} . Therefore T_1 and $\hat{\mu}_1$ are equivalent statistics.

$$E[T_1] = E\left[\sqrt{\frac{\pi}{2}}\frac{1}{n} \sum_{i=1}^n |x_i|\right] = \sqrt{\frac{\pi}{2}}\frac{1}{n} \sum_{i=1}^n E[|x_i|] = \sqrt{\frac{\pi}{2}}\mu_1(\sigma) = \sigma$$

By the empirical substitution principle, $\sqrt{\frac{\pi}{2}}\frac{1}{n} \sum_{i=1}^n |x_i|$ is an estimate of σ based on the first moment.

2.1-15 Hardy-Weinberg with six genotypes.

In a large natural population of plants (*Mimulus guttatus*) there are three possible alleles S, I, and F at one locus resulting in six genotypes labeled SS, II, FF, SI, SF, and IF. Let θ_1, θ_2 , and θ_3 denote the probabilities of S, I, and F, respectively, where $\sum_{j=1}^3 \theta_j = 1$. The Hardy-Weinberg model specifies that the six genotypes have probabilities

Genotype	1	2	3	4	5	6	
Genotype	SS	II	FF	SI	SF	IF	
p_j	$p_1 = \theta_1^2$	$p_2 = \theta_2^2$	$p_3 = \theta_3^2$	$p_4 = 2\theta_1\theta_2$	$p_5 = 2\theta_1\theta_3$	$p_6 = 2\theta_2\theta_3$	$\sum_{j=1}^3 \theta_j = 1$
N_j	N_1	N_2	N_3	N_4	N_5	N_6	$\sum N_j = n$
\hat{p}_j	N_1/n	N_2/n	N_3/n	N_4/n	N_5/n	N_6/n	$\sum \hat{p}_j = 1$

Let N_j be the number of plants of genotype j in a sample of n independent plants, $1 \leq j \leq 6$ and $\hat{p}_j = N_j/n$. Show the frequency plug-in estimates of θ_1, θ_2 , and θ_3

$$\theta_1^2 + \theta_1\theta_2 + \theta_1\theta_3 = \theta_1(\theta_1 + \theta_2 + \theta_3) \stackrel{\sum_{\theta_j=1}}{=} \theta_1 \implies \hat{\theta}_1 = \hat{p}_1 + \frac{1}{2}\hat{p}_4 + \frac{1}{2}\hat{p}_5$$

$$\theta_2^2 + \theta_1\theta_2 + \theta_2\theta_3 = \theta_2(\theta_1 + \theta_2 + \theta_3) \stackrel{\sum_{\theta_j=1}}{=} \theta_2 \implies \hat{\theta}_2 = \hat{p}_2 + \frac{1}{2}\hat{p}_4 + \frac{1}{2}\hat{p}_6$$

$$\theta_3^2 + \theta_1\theta_3 + \theta_2\theta_3 = \theta_3(\theta_1 + \theta_2 + \theta_3) \stackrel{\sum_{\theta_j=1}}{=} \theta_3 \implies \hat{\theta}_3 = \hat{p}_3 + \frac{1}{2}\hat{p}_5 + \frac{1}{2}\hat{p}_6$$

HW4

2.2-12 Let $X_1, \dots, X_n, n \geq 2$, be independently and identically distributed with density

$$f(x, \theta) = \frac{1}{\sigma} \exp\left[-\frac{x-\mu}{\sigma}\right], x \geq \mu, \text{ where } \theta = (\mu, \sigma^2), -\infty < \mu < \infty, \sigma^2 > 0.$$

(a) Find maximum likelihood estimates of μ and σ^2 .

$$L(\theta) = \left(\frac{1}{\sigma}\right)^n \exp\left[-\frac{\sum_{i=1}^n (x_i - \mu)}{\sigma}\right] \mathbf{1}_{\{x_i \geq \mu\}} = \exp\left[-\frac{1}{\sigma} \sum_{i=1}^n x_i + \frac{n\mu}{\sigma} - n \ln \sigma\right] \mathbf{1}_{\{x_i \geq \mu\}}$$

Given σ , $L(\mu)$ is monotone increasing in μ , $\sup L(\mu)$ is equivalent to $\max \mu \leq x_i$, thus $\hat{\mu} = X_{(1)}$ is the maximum likelihood estimator of μ

Given μ

$$l(\theta) = -\frac{1}{\sigma} \sum_{i=1}^n x_i + \frac{n\mu}{\sigma} - n \ln \sigma$$

$$l'(\sigma) = \frac{1}{\sigma^2} \sum_{i=1}^n x_i - \frac{n\mu}{\sigma^2} - \frac{n}{\sigma} \stackrel{!}{=} 0$$

$$\hat{\sigma} = \frac{1}{n} \sum_{i=1}^n x_i - \hat{\mu} = \frac{1}{n} \sum_{i=1}^n x_i - x_{(1)}$$

For $n \geq 2, \frac{1}{n} \sum_{i=1}^n x_i - x_{(1)} > 0, \hat{\sigma} \in (0, \infty)$ then

$\hat{\sigma}^2 = h(\hat{\sigma}) = (\hat{\sigma})^2$ is a 1-1 transformation on $\hat{\sigma} \in (0, \infty)$, $\hat{\sigma}^2$ and $\hat{\sigma}$ are equivalent statistics.

Therefore,

$$\hat{\sigma}^2 = \left(\frac{1}{n} \sum_{i=1}^n x_i - x_{(1)}\right)^2$$

(b) Find the maximum likelihood estimate of $P_\theta[X_1 \geq t]$ for $t > \mu$. Hint: You may use Problem 2.2.16(b).

$$F_X = \int_{-\infty}^x \frac{1}{\sigma} \exp\left[-\frac{x-\mu}{\sigma}\right] \mathbf{1}_{\{x \geq \mu\}} dx = 1 - \exp\left[-\frac{x-\mu}{\sigma}\right]$$

Define $\omega = q(t, \mu, \sigma) = P_\theta[X_1 \geq t] = 1 - F_X(t) = \exp\left[-\frac{1}{\sigma}(t - \mu)\right]$ which is an one to one transformation.

Since $(x_{(1)}, \bar{x} - x_{(1)})$ are MLE of (μ, σ) ,

$\hat{\omega} = q(t, \hat{\mu}, \hat{\sigma}) = \exp\left[-\frac{1}{\hat{\sigma}}(t - \hat{\mu})\right] = \exp\left[-\frac{t - x_{(1)}}{\bar{x} - x_{(1)}}\right]$ is the maximum likelihood estimate of $P_\theta[X_1 \geq t]$ for $t > \mu$

2.2-13 Let X_1, \dots, X_n be a sample from $aU[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$ distribution.

Show that any T such that $X_{(n)} - \frac{1}{2} \leq T \leq X_{(1)} + \frac{1}{2}$ is a maximum likelihood estimate of θ . (We write $U[a, b]$ to make $p(a) = p(b) = (b - a)^{-1}$ rather than 0.)

$$p_\theta(x) = \frac{1}{\theta + \frac{1}{2} - \theta - \frac{1}{2}} \mathbf{1}_{\{\theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2}\}} = \mathbf{1}_{\{\theta - \frac{1}{2} \leq x \leq \theta + \frac{1}{2}\}}$$

$$L(\theta) = \prod_{i=1}^n \mathbf{1}_{\{\theta - \frac{1}{2} \leq x_{(i)} \leq \theta + \frac{1}{2}\}} = \prod_{i=1}^n \mathbf{1}_{\{\theta - \frac{1}{2} \leq x_{(1)}, x_{(n)} \leq \theta + \frac{1}{2}\}} = \mathbf{1}_{\{\theta \leq x_{(1)} + \frac{1}{2}, x_{(n)} - \frac{1}{2} \leq \theta\}} = \mathbf{1}_{\{x_{(n)} - \frac{1}{2} \leq \theta \leq x_{(1)} + \frac{1}{2}\}}$$

To maximize $L(\theta)$ is equivalent to let $X_{(n)} - \frac{1}{2} \leq \theta \leq X_{(1)} + \frac{1}{2}$, which is a MLE of θ .

HW5

2.3-3 Consider the Hardy-Weinberg model with the six genotypes given in Problem 2.1.15.

Let $\Theta = \{(\theta_1, \theta_2) : \theta_1 > 0, \theta_2 > 0, \theta_1 + \theta_2 < 1\}$ and let $\theta_3 = 1 - (\theta_1 + \theta_2)$. In a sample of n independent plants, write $x_i = j$ if the i th plant has genotype j , $1 \leq j \leq 6$. Under what conditions on (x_1, \dots, x_n) does the MLE exist? What is the MLE? Is it unique?

Genotype	1	2	3	4	5	6	
p_j	$p_1 = \theta_1^2$	$p_2 = \theta_2^2$	$p_3 = \theta_3^2$	$p_4 = 2\theta_1\theta_2$	$p_5 = 2\theta_1\theta_3$	$p_6 = 2\theta_2\theta_3$	$\sum_{j=1}^3 \theta_j = 1$
N_j	N_1	N_2	N_3	N_4	N_5	N_6	$\sum N_j = n$

N_j be the number of plants of genotype j in a sample of n independent plants. $\sum N_j = n$. Let

$$t_0 = N_4 + N_5 + N_6$$

$$t_1 = 2N_1 + N_4 + N_5$$

$$t_2 = 2N_2 + N_4 + N_6$$

$$t_3 = 2n - t_1 - t_2 = 2n - (2N_1 + N_4 + N_5) - (2N_2 + N_4 + N_6) = 2N_3 + N_5 + N_6$$

$$\begin{aligned} L(\theta_1, \theta_2, \theta_3 | x_{1:n}) &= (\theta_1^2)^{N_1} (\theta_2^2)^{N_2} (\theta_3^2)^{N_3} (2\theta_1\theta_2)^{N_4} (2\theta_1\theta_3)^{N_5} (2\theta_2\theta_3)^{N_6} \\ &= \exp[(N_4 + N_5 + N_6) \ln 2 + (2N_1 + N_4 + N_5) \ln \theta_1 + (2N_2 + N_4 + N_6) \ln \theta_2 + (2N_3 + N_5 + N_6) \ln \theta_3] \\ &= \exp[t_0 \ln 2 + t_1 \ln \theta_1 + t_2 \ln \theta_2 + t_3 \ln \theta_3] \end{aligned}$$

Substitute θ_3 with $1 - \theta_1 - \theta_2$ and t_3 with $2n - t_1 - t_2$

$$L(\theta_1, \theta_2 | x_{1:n}) = \exp[t_0 \ln 2 + t_1 \ln \theta_1 + t_2 \ln \theta_2 + (2n - t_1 - t_2) \ln(1 - \theta_1 - \theta_2)]$$

$$L(\theta_1, \theta_2 | x_{1:n}) = \underbrace{2^{t_0}}_{h(x)} \exp \left[\underbrace{t_1 \ln \frac{\theta_1}{1 - \theta_1 - \theta_2} + t_2 \ln \frac{\theta_2}{1 - \theta_1 - \theta_2}}_{T(\vec{x}) \cdot \eta(\vec{\theta})} - \underbrace{2n \ln \frac{1}{1 - \theta_1 - \theta_2}}_{A(\eta)} \right]$$

which is a 2-parameter exponential family.

$$l(\theta_1, \theta_2) = t_0 \ln 2 + t_1 \ln \theta_1 + t_2 \ln \theta_2 + t_3 \ln(1 - \theta_1 - \theta_2)$$

$$\begin{aligned} \frac{\partial}{\partial \theta_1} l(\theta_1, \theta_2) &= \frac{t_1}{\theta_1} + \frac{-t_3}{1 - \theta_1 - \theta_2} \stackrel{set}{=} 0 \\ \frac{\partial}{\partial \theta_2} l(\theta_1, \theta_2) &= \frac{t_2}{\theta_2} + \frac{-t_3}{1 - \theta_1 - \theta_2} \stackrel{set}{=} 0 \\ \frac{\partial^2}{\partial \theta_1^2} l(\theta_1, \theta_2) &= \frac{-t_1}{\theta_1^2} + \frac{-t_3}{(1 - \theta_1 - \theta_2)^2} < 0 \\ \frac{\partial^2}{\partial \theta_2^2} l(\theta_1, \theta_2) &= \frac{-t_2}{\theta_2^2} + \frac{-t_3}{(1 - \theta_1 - \theta_2)^2} < 0 \\ \frac{\partial^2}{\partial \theta_1 \partial \theta_2} l(\theta_1, \theta_2) &= \frac{-t_3}{(1 - \theta_1 - \theta_2)^2} \end{aligned}$$

$$\begin{aligned}
\left[\frac{\partial^2}{\partial\theta_1^2}l(\theta_1, \theta_2)\right]\left[\frac{\partial^2}{\partial\theta_2^2}l(\theta_1, \theta_2)\right] &= \left(\frac{-t_1}{\theta_1^2} + \frac{-t_3}{(1-\theta_1-\theta_2)^2}\right) \cdot \left(\frac{-t_2}{\theta_2^2} + \frac{-t_3}{(1-\theta_1-\theta_2)^2}\right) \\
&= \frac{t_1 t_2}{\theta_1^2 \theta_2^2} + \left(\frac{t_1}{\theta_1^2} + \frac{t_2}{\theta_2^2}\right) \frac{t_3}{(1-\theta_1-\theta_2)^2} + \left[\frac{-t_3}{(1-\theta_1-\theta_2)^2}\right]^2 \\
&> \left[\frac{-t_3}{(1-\theta_1-\theta_2)^2}\right]^2 \\
&= \left[\frac{\partial^2}{\partial\theta_1 \partial\theta_2}l(\theta_1, \theta_2)\right]^2
\end{aligned}$$

Therefore, the likelihood function is strictly concave and the unique MLEs exist. The MLE solutions are

$$\begin{cases} \hat{\theta}_1 = \frac{t_1}{2n} = \frac{2N_1+N_4+N_5}{2n} \\ \hat{\theta}_2 = \frac{t_2}{2n} = \frac{2N_2+N_4+N_6}{2n} \\ \hat{\theta}_3 = \frac{t_3}{2n} = \frac{2N_3+N_5+N_6}{2n} \end{cases}$$

2.3-12 Let X_1, \dots, X_n be i.i.d. $\frac{1}{\sigma}f_0(\frac{x-\mu}{\sigma})$, $\sigma > 0, \mu \in R$, and assume for $w \equiv -\log f_0$ that $w'' > 0$ so that w is strictly convex, $w(\pm\infty) = \infty$.

- (a) Show that, if $n \geq 2$, the likelihood equations $\sum_{i=1}^n w'(\frac{X_i-\mu}{\sigma}) = 0$; $\sum_{i=1}^n [\frac{(X_i-\mu)}{\sigma} w'(\frac{X_i-\mu}{\sigma}) - 1] = 0$ have a unique solution $(\hat{\mu}, \hat{\sigma})$.

$$\begin{aligned}
L(\mu, \sigma) &= \prod_{i=1}^n \frac{1}{\sigma} f_0\left(\frac{x_i - \mu}{\sigma}\right) \\
l(\mu, \sigma) &= n \ln\left(\frac{1}{\sigma}\right) + \sum_{i=1}^n \ln f_0\left(\frac{x_i - \mu}{\sigma}\right) = n \ln\left(\frac{1}{\sigma}\right) - \sum_{i=1}^n w\left(\frac{x_i - \mu}{\sigma}\right)
\end{aligned}$$

By *Hint (b)* Reparametrize by $a = \frac{1}{\sigma}$, $b = \frac{\mu}{\sigma}$ and consider varying a , b successively.

By *Hint (a)* The function $D(a, b) = \sum_{i=1}^n w(ax_i - b) - n \log a$ is strictly convex in (a, b) and $\lim_{(a,b) \rightarrow (a_0, b_0)} D(a, b) = \infty$ if either $a_0 = 0$ or ∞ or $b_0 = \pm\infty$,

By *Hint (ii)* If $\frac{\partial^2 D}{\partial a^2} > 0$, $\frac{\partial^2 D}{\partial b^2} > 0$ and $\frac{\partial^2 D}{\partial a^2} \frac{\partial^2 D}{\partial b^2} > (\frac{\partial^2 D}{\partial a \partial b})^2$, then D is strictly convex.

Check The function $D(a, b) = \sum_{i=1}^n w(ax_i - b) - n \log a$

$$\begin{aligned}
\frac{\partial D}{\partial a} &= -\frac{n}{a} + \sum_{i=1}^n w'(ax_i - b)(x_i) \\
\frac{\partial^2 D}{\partial a^2} &= \frac{n}{a^2} + \sum_{i=1}^n w''(ax_i - b)(x_i^2) > 0 \\
\frac{\partial D}{\partial b} &= \sum_{i=1}^n w'(ax_i - b)(-1) \\
\frac{\partial^2 D}{\partial b^2} &= \sum_{i=1}^n w''(ax_i - b) > 0 \\
\frac{\partial^2 D}{\partial a \partial b} &= \sum_{i=1}^n w''(ax_i - b)(-x_i)
\end{aligned}$$

$$\begin{aligned}
\left[\frac{\partial^2 D}{\partial a^2}\right]\left[\frac{\partial^2 D}{\partial b^2}\right] &= \left[\frac{n}{a^2} + \sum_{i=1}^n w''(ax_i - b)(x_i^2)\right] \cdot \sum_{i=1}^n w''(ax_i - b) \\
&= \frac{n}{a^2} \sum_{i=1}^n w''(ax_i - b) + \left[\sum_{i=1}^n w''(ax_i - b)\right]^2 (x_i^2) \\
&> \left[\sum_{i=1}^n w''(ax_i - b)(x_i)\right]^2 = \left[\frac{\partial D^2}{\partial a \partial b}\right]^2
\end{aligned}$$

The function $D(a, b)$ is strictly convex in (a, b) and $\lim_{(a,b) \rightarrow (a_0, b_0)} D(a, b) = \infty$, if either $a_0 = 0$ or ∞ or $b_0 = \pm\infty$

$$l(\mu, \sigma) = n \ln(a) - \sum_{i=1}^n \omega(ax_i - b) = -D(a, b)$$

Therefore, $l(\mu, \sigma)$ is strictly concave in (μ, σ) By *Hint (i)* If a strictly convex function has a minimum, it is unique. The equations have a unique solution $(\hat{\mu}, \hat{\sigma})$ to get a maximum for $L(\mu, \sigma)$

$$\begin{aligned}
\frac{\partial}{\partial \mu} l(\mu, \sigma) &= \sum_{i=1}^n \omega' \left(\frac{x_i - \mu}{\sigma} \right) \frac{-1}{\sigma} \stackrel{set}{=} 0 \\
\frac{\partial}{\partial \sigma} l(\mu, \sigma) &= \frac{-n}{\sigma} + \sum_{i=1}^n \omega' \left(\frac{x_i - \mu}{\sigma} \right) \cdot \frac{x_i - \mu}{\sigma^2} = \frac{1}{\sigma} \sum_{i=1}^n \left[\frac{(x_i - \mu)}{\sigma} w' \left(\frac{x_i - \mu}{\sigma} \right) - 1 \right] \stackrel{set}{=} 0
\end{aligned}$$

$$\begin{aligned}
\sum_{i=1}^n w' \left(\frac{X_i - \hat{\mu}}{\hat{\sigma}} \right) &= 0 \\
\sum_{i=1}^n \left[\frac{(X_i - \hat{\mu})}{\hat{\sigma}} w' \left(\frac{X_i - \hat{\mu}}{\hat{\sigma}} \right) - 1 \right] &= 0
\end{aligned}$$

(b) Give an algorithm such that starting at $\hat{\mu}^0 = 0, \hat{\sigma}^0 = 1, \hat{\mu}^{(i)} \rightarrow \hat{\mu}, \hat{\sigma}^{(i)} \rightarrow \hat{\sigma}$.

Using **Coordinate-Ascent Algorithm**, get the unique solution of $(\hat{\mu}^{(i)}, \hat{\sigma}^{(i)})$ by solving

$$\begin{aligned}
\sum_{i=1}^n w' \left(\frac{X_i - \hat{\mu}^{(1)}}{\hat{\sigma}^{(0)}} \right) &= 0 && \text{plugin } \hat{\sigma}^0, \text{ get } \hat{\mu}^{(1)} \\
\sum_{i=1}^n \left[\frac{(X_i - \hat{\mu}^{(1)})}{\hat{\sigma}^{(1)}} w' \left(\frac{X_i - \hat{\mu}^{(1)}}{\hat{\sigma}^{(1)}} \right) - 1 \right] &= 0 && \text{plugin } \hat{\mu}^{(1)}, \text{ get } \hat{\sigma}^{(1)} \\
&\dots && \\
\sum_{i=1}^n w' \left(\frac{X_i - \hat{\mu}^{(i)}}{\hat{\sigma}^{(i-1)}} \right) &= 0 && \text{plugin } \hat{\sigma}^{(i-1)}, \text{ get } \hat{\mu}^{(i)} \\
\sum_{i=1}^n \left[\frac{(X_i - \hat{\mu}^{(i)})}{\hat{\sigma}^{(i)}} w' \left(\frac{X_i - \hat{\mu}^{(i)}}{\hat{\sigma}^{(i)}} \right) - 1 \right] &= 0 && \text{plugin } \hat{\mu}^{(i)}, \text{ get } \hat{\sigma}^{(i)}
\end{aligned}$$

Using **Newton-Raphson Method**, get the unique solution of $(\hat{\mu}^{(i)}, \hat{\sigma}^{(i)})$ by solving

$$\begin{aligned}\hat{\mu}^{(i)} &= \hat{\mu}^{(i-1)} - \frac{l(\mu, \sigma)}{\frac{\partial}{\partial \mu} l(\mu, \sigma)} = \hat{\mu}^{(i-1)} - \frac{n \ln(\frac{1}{\sigma^{(i-1)}}) - \sum_{i=1}^n \omega(\frac{x_i - \hat{\mu}^{(i-1)}}{\hat{\sigma}^{(i-1)}})}{\sum_{i=1}^n w'(\frac{X_i - \hat{\mu}^{(i-1)}}{\hat{\sigma}^{(i-1)}})} \quad \text{plugin } \hat{\mu}^{(i-1)} \text{ and } \hat{\sigma}^{(i-1)}, \text{ get } \hat{\mu}^i \\ \hat{\sigma}^{(i)} &= \hat{\sigma}^{(i-1)} - \frac{l(\mu, \sigma)}{\frac{\partial}{\partial \sigma} l(\mu, \sigma)} = \hat{\sigma}^{(i-1)} - \frac{n \ln(\frac{1}{\sigma^{(i-1)}}) - \sum_{i=1}^n \omega(\frac{x_i - \hat{\mu}^{(i-1)}}{\hat{\sigma}^{(i-1)}})}{\sum_{i=1}^n [\frac{(X_i - \hat{\mu}^{(i-1)})}{\hat{\sigma}^{(i-1)}} w'(\frac{X_i - \hat{\mu}^{(i-1)}}{\hat{\sigma}^{(i-1)}}) - 1]} \quad \text{plugin } \hat{\mu}^{(i-1)} \text{ and } \hat{\sigma}^{(i-1)}, \text{ get } \hat{\sigma}^i\end{aligned}$$

By **Theorem 2.4.2**, f_0 is the canonical exponential family generated by (T, h) , the natural parameter space ε is open and the family is of rank k . Set $t_0 = T(x)$, $t_0 \in C_T^0$, $\hat{\eta}^{(i)} = (\hat{\mu}^{(i)}, \hat{\sigma}^{(i)}) \rightarrow \hat{\eta} = (\hat{\mu}, \hat{\sigma})$ as $i \rightarrow \infty$

- (c) Show that for the logistic distribution $F_0(x) = [1 + \exp\{-x\}]^{-1}$, w is strictly convex and give the likelihood equations for μ and σ . (See Example 2.4.3.)

$$F_0'(\frac{x - \mu}{\sigma}) = \frac{1}{\sigma} f_0(\frac{x - \mu}{\sigma}) = \frac{1}{\sigma} \frac{\exp[-\frac{x - \mu}{\sigma}]}{(1 + \exp[-\frac{x - \mu}{\sigma}])^2}$$

$$\omega = -\ln f_0(x) = 2 \ln(1 + \exp[-\frac{x - \mu}{\sigma}]) + \frac{x - \mu}{\sigma}$$

Check w is strictly convex.

$$\begin{aligned}\frac{\partial \omega}{\partial x} &= \frac{1}{\sigma} - \frac{2 \exp[-\frac{x - \mu}{\sigma}]}{\sigma(1 + \exp[-\frac{x - \mu}{\sigma}])} = \frac{1 - \exp[-\frac{x - \mu}{\sigma}]}{\sigma(1 + \exp[-\frac{x - \mu}{\sigma}])} \\ \frac{\partial^2 \omega}{\partial x^2} &= \frac{2 \exp[-\frac{x - \mu}{\sigma}]}{\sigma^2(1 + \exp[-\frac{x - \mu}{\sigma}])^2} = \frac{2}{\sigma^2} f_0(x) > 0\end{aligned}$$

Therefore, w is strictly convex. By part (a), $l(\mu, \sigma) = -\sum_{i=1}^n \omega$ is strictly concave in (μ, σ) . The likelihood equations for μ and σ are

$$\begin{aligned}\sum_{i=1}^n w'(\frac{X_i - \hat{\mu}}{\hat{\sigma}}) &= \sum_{i=1}^n \frac{1 - \exp[-\frac{x_i - \hat{\mu}}{\hat{\sigma}}]}{\hat{\sigma}(1 + \exp[-\frac{x_i - \hat{\mu}}{\hat{\sigma}}])} = 0 \\ \sum_{i=1}^n [\frac{(X_i - \hat{\mu})}{\hat{\sigma}} w'(\frac{X_i - \hat{\mu}}{\hat{\sigma}}) - 1] &= \sum_{i=1}^n \left[\frac{(x_i - \hat{\mu})(1 - \exp[-\frac{x_i - \hat{\mu}}{\hat{\sigma}}])}{\hat{\sigma}^2(1 + \exp[-\frac{x_i - \hat{\mu}}{\hat{\sigma}}])} - 1 \right] = 0\end{aligned}$$

HW6

2.4-1 EM for bivariate data.

- (a) In the bivariate normal Example 2.4.6, complete the E-step by finding $E(Z_i|Y_i)$, $E(Z_i^2|Y_i)$ and $E(Z_i Y_i|Y_i)$.

From Example 2.4.6. Let $(Z_1, Y_1), \dots, (Z_n, Y_n)$ be i.i.d. as (Z, Y) , where $(Z, Y) \sim N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$. Suppose that some of the Z_i and some of the Y_i are missing as follows: For $1 \leq i \leq n_1$ we observe both Z_i and Y_i , for $n_1 + 1 \leq i \leq n_2$, we observe only Z_i , and for $n_2 + 1 \leq i \leq n$, we observe only Y_i . In this case a set of sufficient statistics is

$$\begin{aligned} T_1 = \bar{Z} & \quad T_3 = n^{-1} \sum_{i=1}^n Z_i^2 \\ T_2 = \bar{Y} & \quad T_4 = n^{-1} \sum_{i=1}^n Y_i^2 \quad T_5 = n^{-1} \sum_{i=1}^n Z_i Y_i \end{aligned}$$

The observed data are $S = \{(Z_i, Y_i) : 1 \leq i \leq n_1\} \cup \{Z_i : n_1 + 1 \leq i \leq n_2\} \cup \{Y_i : n_2 + 1 \leq i \leq n\}$

To compute $E\theta(\mathbf{T}|S = s)$, where $\theta = (\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$, we note that for the cases with Z_i and/or Y_i observed, the conditional expected values equal their observed values.

By **Theorem B.4.2** If (X, Y) has a nondegenerate $N(\mu_1, \mu_2, \sigma_1^2, \sigma_2^2, \rho)$ distribution, then the conditional distribution of Y given $X = x$ is $N\left(\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(x - \mu_1), \sigma_2^2(1 - \rho^2)\right)$.

Use the properties of the bivariate normal distribution to conclude

$$\begin{aligned} E_\theta(Y_i|Z_i) &= \mu_2 + \rho \frac{\sigma_2}{\sigma_1}(Z_i - \mu_1) \\ E_\theta(Y_i^2|Z_i) &= [\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(Z_i - \mu_1)]^2 + (1 - \rho^2)\sigma_2^2 \\ E_\theta(Z_i Y_i|Z_i) &= [\mu_2 + \rho \frac{\sigma_2}{\sigma_1}(Z_i - \mu_1)]Z_i \end{aligned}$$

The conditional distribution of $(Z|Y) \sim N\left(\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(Y - \mu_2), \sigma_1^2(1 - \rho^2)\right)$ with the corresponding Z on Y regression equations.

$$\begin{aligned} E_\theta(Z_i|Y_i) &= \mu_1 + \rho \frac{\sigma_1}{\sigma_2}(Y_i - \mu_2) \\ E_\theta(Z_i^2|Y_i) &= [\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(Y_i - \mu_2)]^2 + (1 - \rho^2)\sigma_1^2 \\ E_\theta(Z_i Y_i|Y_i) &= [\mu_1 + \rho \frac{\sigma_1}{\sigma_2}(Y_i - \mu_2)]Y_i \end{aligned}$$

The E-step involves imputing missing values in the context of Example 2.4.6. The EM algorithm is a form of multiple imputation.

- (b) In Example 2.4.6, verify the M-step by showing that $E_\theta \mathbf{T} = (\mu_1, \mu_2, \sigma_1^2 + \mu_1^2, \sigma_2^2 + \mu_2^2, \rho\sigma_1\sigma_2 + \mu_1\mu_2)$.

$$E[T_1] = E[\bar{Z}] = n^{-1} \sum_{i=1}^n E[Z_i] = \mu_1$$

$$E[T_2] = E[\bar{Y}] = n^{-1} \sum_{i=1}^n E[Y_i] = \mu_2$$

$$E[T_3] = E[n^{-1} \sum_{i=1}^n Z_i^2] = n^{-1} \sum_{i=1}^n (V[Z_i] + E[Z_i]^2) = \sigma_1^2 + \mu_1^2$$

$$E[T_4] = E[n^{-1} \sum_{i=1}^n Y_i^2] = n^{-1} \sum_{i=1}^n (V[Y_i] + E[Y_i]^2) = \sigma_2^2 + \mu_2^2$$

$$E[T_5] = E[n^{-1} \sum_{i=1}^n Z_i Y_i] = n^{-1} \sum_{i=1}^n (Cov[Z_i Y_i] + E[Z_i] E[Y_i]) = n^{-1} \sum_{i=1}^n (\rho \sqrt{V[Z_i] V[Y_i]} + \mu_1 \mu_2) = \rho \sigma_1 \sigma_2 + \mu_1 \mu_2$$

$$\dot{A}(\theta) = E_{\theta} \mathbf{T} = (\mu_1, \mu_2, \sigma_1^2 + \mu_1^2, \sigma_2^2 + \mu_2^2, \sigma_1 \sigma_2 \rho + \mu_1 \mu_2)$$

We take $\hat{\theta}_{old} = \hat{\theta}_{MOM}$, where $\hat{\theta}_{MOM}$ is the method of moment estimates $(\hat{\mu}_1, \hat{\mu}_2, \hat{\sigma}_1^2, \hat{\sigma}_2^2, r)$ (Problem 2.1.8) of θ based on the observed data. The M-step produces

$$\begin{aligned} \hat{\mu}_{1,new} &= T_1(\hat{\theta}_{old}) \\ \hat{\mu}_{2,new} &= T_2(\hat{\theta}_{old}) \\ \hat{\sigma}_{1,new}^2 &= T_3(\hat{\theta}_{old}) - \hat{T}_1^2 \\ \hat{\sigma}_{2,new}^2 &= T_4(\hat{\theta}_{old}) - \hat{T}_2^2 \\ \hat{\rho}_{new} &= \frac{T_5(\hat{\theta}_{old}) - \hat{T}_1 \hat{T}_2}{\{[T_3(\hat{\theta}_{old}) - \hat{T}_1][T_4(\hat{\theta}_{old}) - \hat{T}_2]\}^{\frac{1}{2}}} \end{aligned}$$

where $T_j(\theta)$ denotes T_j with missing values replaced by the values computed in the E-step and $\hat{T}_j = T_j(\hat{\theta}_{old})$, $j = 1, 2$. Now the process is repeated with $\hat{\theta}_{MOM}$ replaced by $\hat{\theta}_{new}$.

2.4-6 Consider a genetic trait that is directly unobservable but will cause a disease among a certain proportion of the individuals that have it.

For families in which one member has the disease, it is desired to estimate the proportion θ that has the genetic trait. Suppose that in a family of n members in which one has the disease (and, thus, also the trait), X is the number of members who have the trait. Because it is known that $X \geq 1$, the model often used for X is that it has the conditional distribution of a $\mathcal{B}(n, \theta)$ variable, $\theta \in [0, 1]$, given $X \geq 1$.

(a) Show that $P(X = x | X \geq 1) = \frac{\binom{n}{x} \theta^x (1-\theta)^{n-x}}{1 - (1-\theta)^n}$, $x = 1, \dots, n$, and that the MLE exists and is unique.

$$\begin{aligned}
P(X = x|X \geq 1) &= \frac{P(X = x, X \geq 1)}{P(X \geq 1)} = \frac{P(X \geq 1|X = x)P(X = x)}{1 - P(X = 0)} \\
&= \frac{\binom{n}{x}\theta^x(1-\theta)^{n-x}}{1 - \binom{n}{0}\theta^0(1-\theta)^{n-0}} \mathbf{1}_{\{x \in \mathbf{N}^+\}} \\
&= \underbrace{\binom{n}{x} \mathbf{1}_{\{x \in \mathbf{N}^+\}}}_{h(x)} \exp \left(\underbrace{x \ln \left[\frac{\theta}{1-\theta} \right]}_{\eta} - \underbrace{\ln \frac{1 - (1-\theta)^n}{(1-\theta)^n}}_{\eta(\theta)} \right) \\
\eta = \ln \left[\frac{\theta}{1-\theta} \right] &\implies \frac{\theta}{1-\theta} = e^\eta; \quad 1-\theta = \frac{1}{e^\eta + 1}; \quad \theta = \frac{e^\eta}{e^\eta + 1} \\
\eta(\theta) &= \ln \frac{1 - (1-\theta)^n}{(1-\theta)^n} = \ln[(e^\eta + 1)^n - 1]
\end{aligned}$$

By **Theorem 2.3.1** This is a canonical exponential family generated by (T, h) .

$\eta(\theta)$ is finite for $\eta > -\infty$

(i) The natural parameter space, $\varepsilon = \mathbb{R}$, is open.

(ii) The family is of rank 1.

x is the observed data vector and set $t_0 = T(x) = x$.

$$P[c^T T(X) > c^T t_0] = P[cX > cx] > 0, \quad \forall c \neq 0, x \in \mathbb{R}^{n-1}$$

Thus, the MLE $\hat{\eta}$ exists, is unique, and is a solution to the equation

$$A(\eta) = E_\eta(T(X)) = t_0$$

(b) Use (2.4.3) to show that the Newton-Raphson algorithm gives $\hat{\theta}_1 = \tilde{\theta} - \frac{\tilde{\theta}(1-\tilde{\theta})[1-(1-\tilde{\theta})^n]\{x-n\tilde{\theta}-x(1-\tilde{\theta})^n\}}{n\tilde{\theta}^2(1-\tilde{\theta})^n[n-1+(1-\tilde{\theta})^n]-[1-(1-\tilde{\theta})^n]^2[(1-2\tilde{\theta})x+n\tilde{\theta}^2]}$, where $\tilde{\theta} = \hat{\theta}_{old}$ and $\hat{\theta}_1 = \hat{\theta}_{new}$, as the first approximation to the maximum likelihood estimate of θ .

$$\begin{aligned}
l(\theta) &= \ln \binom{n}{x} + \ln \mathbf{1}_{\{x \in \mathbf{N}^+\}} + x \ln \theta - x \ln(1-\theta) - \ln[1 - (1-\theta)^n] + n \ln(1-\theta) \\
\dot{l}(\theta) &= \frac{x}{\theta} - \frac{x}{1-\theta} + \frac{n}{1-\theta}(-1) - \frac{1}{1-(1-\theta)^n}(-n)(1-\theta)^{n-1}(-1) = \frac{x}{\theta(1-\theta)} - \frac{n(1-\theta)^{n-1}}{1-(1-\theta)^n} - \frac{n}{1-\theta} \\
\ddot{l}(\theta) &= \frac{-x}{\theta^2} + \frac{x}{(1-\theta)^2} - \frac{n}{(1-\theta)^2} + \frac{n(n-2)(1-\theta)^{n-2}[1-(1-\theta)^n] + n(1-\theta)^{n-1}n(1-\theta)^{n-1}}{[1-(1-\theta)^n]^2}
\end{aligned}$$

$$\text{By (2.4.3) } \hat{\theta}_{new} = \hat{\theta}_{old} - \ddot{l}^{-1}(\hat{\theta}_{old})\dot{l}(\hat{\theta}_{old})$$

$$\begin{aligned}
\hat{\theta}_1 &= \tilde{\theta} - \frac{\frac{x}{\tilde{\theta}(1-\tilde{\theta})} - \frac{n}{1-\tilde{\theta}} - \frac{n(1-\tilde{\theta})^{n-1}}{1-(1-\tilde{\theta})^n}}{\frac{-x}{\tilde{\theta}^2} + \frac{x}{(1-\tilde{\theta})^2} + \frac{n(n-1)(1-\tilde{\theta})^{n-2}[1-(1-\tilde{\theta})^n] + n(1-\tilde{\theta})^{n-1}n(1-\tilde{\theta})^{n-1}}{[1-(1-\tilde{\theta})^n]^2} - \frac{n}{(1-\tilde{\theta})^2}} = \\
&= \tilde{\theta} - \frac{\tilde{\theta}(1-\tilde{\theta})[1-(1-\tilde{\theta})^n] \{ [1-(1-\tilde{\theta})^n]x - n\tilde{\theta}[1-(1-\tilde{\theta})^n] - \tilde{\theta}(1-\tilde{\theta})n(1-\tilde{\theta})^{n-1} \}}{[1-(1-\tilde{\theta})^n]^2 \{ -(1-\tilde{\theta})^2x + \tilde{\theta}^2x - n\tilde{\theta}^2 \} + \tilde{\theta}^2(1-\tilde{\theta})^2 \{ n(n-1)(1-\tilde{\theta})^{n-2}[1-(1-\tilde{\theta})^n] + n(1-\tilde{\theta})^{n-1}n(1-\tilde{\theta})^{n-1} \}} \\
&= \tilde{\theta} - \frac{\tilde{\theta}(1-\tilde{\theta})[1-(1-\tilde{\theta})^n] \{ [1-(1-\tilde{\theta})^n]x - n\tilde{\theta}[1-(1-\tilde{\theta})^n] - n\tilde{\theta}(1-\tilde{\theta})^n \}}{[1-(1-\tilde{\theta})^n]^2 \{ (2\tilde{\theta}-1)x - n\tilde{\theta}^2 \} + n\tilde{\theta}^2(1-\tilde{\theta})^n \{ (n-1)[1-(1-\tilde{\theta})^n] + n(1-\tilde{\theta})^n \}} \\
&= \tilde{\theta} - \frac{\tilde{\theta}(1-\tilde{\theta})[1-(1-\tilde{\theta})^n] \{ [1-(1-\tilde{\theta})^n]x - n\tilde{\theta} \}}{[1-(1-\tilde{\theta})^n]^2 \{ (2\tilde{\theta}-1)x - n\tilde{\theta}^2 \} + n\tilde{\theta}^2(1-\tilde{\theta})^n \{ n-1+(1-\tilde{\theta})^n \}} \\
&= \tilde{\theta} - \frac{\tilde{\theta}(1-\tilde{\theta})[1-(1-\tilde{\theta})^n] \{ x - n\tilde{\theta} - x(1-\tilde{\theta})^n \}}{n\tilde{\theta}^2(1-\tilde{\theta})^n \{ n-1+(1-\tilde{\theta})^n \} - [1-(1-\tilde{\theta})^n]^2 \{ (1-2\tilde{\theta})x + n\tilde{\theta}^2 \}}
\end{aligned}$$

(c) If $n = 5, x = 2$, find $\hat{\theta}_1$ of (b) above using $\theta = x/n$ as a preliminary estimate.

$$\tilde{\theta} = \frac{2}{5}$$

$$\begin{aligned}
\hat{\theta}_1 &= \frac{2}{5} - \frac{\frac{2}{5}(1-\frac{2}{5})[1-(1-\frac{2}{5})^5] \{ 2-5(\frac{2}{5})-2(1-\frac{2}{5})^5 \}}{5(\frac{2}{5})^2(1-\frac{2}{5})^5 \{ 5-1+(1-\frac{2}{5})^5 \} - [1-(1-\frac{2}{5})^5]^2 \{ (1-2 \cdot \frac{2}{5})2+5(\frac{2}{5})^2 \}} \\
&= \frac{2}{5} - \frac{\frac{2}{5}(\frac{3}{5})[1-(\frac{3}{5})^5] \{ 2-2-2(\frac{3}{5})^5 \}}{\frac{4}{5}(\frac{3}{5})^5 \{ 4+(\frac{3}{5})^5 \} - [1-(\frac{3}{5})^5]^2 \{ (1-\frac{4}{5})2+\frac{4}{5} \}} \\
&= \frac{2}{5} - \frac{\frac{6}{25}[1-\frac{243}{3125}]\{-2(\frac{243}{3125})\}}{\frac{4}{5}(\frac{243}{3125}) \{ 4+\frac{243}{3125} \} - [1-\frac{243}{3125}]^2 \frac{6}{5}} = \frac{2}{5} - \frac{-\frac{12}{25}[\frac{2882}{3125}](\frac{243}{3125})}{\frac{4}{5}(\frac{243}{3125}) \{ \frac{12743}{3125} \} - [\frac{2882}{3125}]^2 \frac{6}{5}} \\
&= \frac{2}{5} - \frac{-12[700326]}{20(3096549) - [8305924]30} = \frac{2}{5} - \frac{-8403912}{-187246740} = 0.4 - 0.04488149 = 0.3551185
\end{aligned}$$