

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 $h(x) = 1$, $a(\theta) = (\theta_2 - \theta_1)^{-1}$

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

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

$$h'(x) = 1$$

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

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.

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

Thus, $\sum_{i=1}^n X_i$ is a sufficient statistic for θ . $P_\theta(\sum_{i=1}^n X_i = x)$ is a one-parameter exponential family.

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

In part (b), $P_\theta(\sum_{i=1}^n X_i = k) = (1 - \theta)^k \theta^n$.

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

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

$$\frac{d^k}{d\omega^k} (1 - \omega)^{-n} = (-n-k+1)(-1) \cdots (n+1)n(1 - \omega)^{-n-k+1} = \left[\prod_{i=0}^{k-1} (n+i) \right] (1 - \omega)^{-n-k+1}$$

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

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

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.

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

(b) The gamma family.

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

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 .

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

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

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

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 $0 < \theta < 1$.

- Show that $T_3 = N_1/n + N_2/2n$ is a frequency substitution estimate of θ .
- Using the estimate of (a), what is a frequency substitution estimate of the odds ratio $\frac{\theta}{1-\theta}$?
- 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 .

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

- Find an estimate of σ^2 based on the second moment.
- Construct an estimate of σ using the estimate of part (a) and the equation $\sigma = \sqrt{\sigma^2}$.
- Use the empirical substitution principle to construct an estimate of σ using the relation $E(|X_1|) = \sigma\sqrt{2/\pi}$.

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
Probability	θ_1^2	θ_2^2	θ_3^2	$2\theta_1\theta_2$	$2\theta_1\theta_3$	$2\theta_2\theta_3$

Let N_j be the number of plants of genotype j in a sample of n independent plants, $1 \leq j \leq 6$ and let $\hat{p}_j = N_j/n$. Show that

$$\hat{\theta}_1 = \hat{p}_1 + 12\hat{p}_4 + 12\hat{p}_5$$

$$\hat{\theta}_2 = \hat{p}_2 + 12\hat{p}_4 + 12\hat{p}_6$$

$$\hat{\theta}_3 = \hat{p}_3 + 12\hat{p}_5 + 12\hat{p}_6$$

are frequency plug-in estimates of θ_1, θ_2 , and θ_3 .

HW4

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

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

- Find maximum likelihood estimates of μ and σ^2 .
- Find the maximum likelihood estimate of $P_\theta[X_1 \geq t] \text{ for } t > \mu$. Hint: You may use Problem 2.2.16(b).

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.)

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?

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 $(\hat{\mu}, \hat{\sigma})$.

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

(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.)

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 $b_0 = \pm\infty$.

(b) Reparametrize by $a = \frac{1}{\sigma}$, $b = \frac{\mu}{\sigma}$ and consider varying a, b successively. Note: You may use without proof (see Appendix B.9).

(c) If a strictly convex function has a minimum, it is unique.

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

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 $E(Z_i Y_i|Y_i)$.

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

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

(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 $\tilde{\theta} = \hat{\theta}_{\text{old}}$ and $\hat{\theta}_1 = \hat{\theta}_{\text{new}}$, as the first approximation to the maximum likelihood estimate of θ .

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