Problem 1. Inversion Theorems

Show the one-to-one relationship between the probability distribution function and the characteristic function. Express the probability distribution function as a function of the characteristic function.

Problem 2.

Explain why the moment generation function does not exist when a certain moment is absent. Provide an explanation for the absence of a moment in the context of the logistic distribution.

Problem 3.

Logistic distribution 有對稱性。
説明：其他 model 沒有對稱性

Problem 4.

(b.1) Log-linear model:

$$P(Y=y)=C\left(\theta_1,\theta_2\right)\exp\left(Y^T\theta_1+W^T\theta_2\right),$$

where

- $W=\left(Y_1Y_2,Y_1Y_3,\cdots,Y_{m-1}Y_m,\cdots,Y_1Y_2\cdots Y_m\right)^T$,
- $\theta_1=\left(\theta_1^{(1)},\cdots,\theta_m^{(1)}\right)$
- $\theta_2=\left(\theta_{12}^{(2)},\cdots,\theta_{m-1m}^{(2)},\cdots,\theta_{12\cdots m}^{(m)}\right)$
- $C(\theta_1,\theta_2)$ is a function of θ_1 and θ_2 that normalizes the p.d.f. to integrate to one.

Transformation:

$$\left(\theta_1,\theta_2\right)\to\left(\mu,\theta_2\right),\mu=\left(\mu_1,\ldots,\mu_m\right)\triangleq\mu\left(\theta_1,\theta_2\right).$$

Model assumption:

$$\operatorname{logit}\left(\mu_j\right)=X_j^T\beta.$$

The score equation for β under this parameterization takes the GEE form:

$$\left(\frac{\partial\mu}{\partial\beta}\right)^T[V(Y)]^{-1}(Y-\mu)=0,$$

where

$$\frac{\partial\mu}{\partial\beta}=\left(\frac{\partial\mu_1}{\partial\beta},\ldots,\frac{\partial\mu_m}{\partial\beta}\right)^T.$$

Remark:
The conditional odds ratios are not easily interpreted when the association among responses is itself a focus of the study.
Properties:
1. From

$$M_Y(t)=E\left[e^{t^TY}\right]=\sum_y C\left(\theta_1,\theta_2\right)\exp\left(y^T\left(t+\theta_1\right)+w^T\theta_2\right)=\frac{C\left(\theta_1,\theta_2\right)}{C\left(\theta_1+t,\theta_2\right)},$$

one has

- $$\mu=\left.\frac{\partial M_Y(t)}{\partial t}\right|_{t=0}=-\frac{\frac{\partial}{\partial\theta_1}C\left(\theta_1,\theta_2\right)}{C\left(\theta_1,\theta_2\right)},$$
- $$E\left[YY^T\right]=\left.\frac{\partial^2M_Y(t)}{\partial t\partial t^T}\right|_{t=0}=-\frac{\frac{\partial^2}{\partial\theta_1\partial\theta_1^T}C\left(\theta_1,\theta_2\right)}{C\left(\theta_1,\theta_2\right)}+2\mu\mu^T,$$
- $$V(Y)=-\frac{\frac{\partial^2}{\partial\theta_1\partial\theta_1^T}C\left(\theta_1,\theta_2\right)}{C\left(\theta_1,\theta_2\right)}+\mu\mu^T.$$

2. Let

$$l(\theta_1, \theta_2) = \ln P(Y = y) = \ln\{C(\theta_1, \theta_2) + (Y^T\theta_1 + W^T\theta_2)\}.$$

We can derive that

$$\frac{\partial l\left(\theta_1, \theta_2\right)}{\partial \theta_1} = \frac{\frac{\partial}{\partial \theta_1} C\left(\theta_1, \theta_2\right)}{C\left(\theta_1, \theta_2\right)} + Y = (Y - \mu),$$

and hence,

$$\frac{\partial l\left(\theta_1, \theta_2\right)}{\partial \beta} = \left(\frac{\partial \mu}{\partial \beta}\right)^T \frac{\partial \theta_1}{\partial \mu}\left(\frac{\partial l\left(\theta_1, \theta_2\right)}{\partial \theta_1}\right) = \left(\frac{\partial \mu}{\partial \beta}\right)^T\left(\frac{\partial \mu}{\partial \theta_1}\right)^{-1}(Y-\mu) = \left(\frac{\partial \mu}{\partial \beta}\right)^T [V(Y)]^{-1}(Y-\mu).$$

Problem 5. Show orthonormal

Let $P_{[1]}(Y = y) = \prod_{j=1}^{m} \mu_j^{y_j}\left(1 - \mu_j\right)^{1-y_j}, g(y) = P(Y = y)/P_{[1]}(Y = y)$, and V be a vector space of real-valued functions f on Y_1 (2^m possible values of y).
Here, V is regarded as an inner-product space with

$$< f_1, f_2 >\triangleq E_{P_{[1]}}\left[f_1 f_2\right] = \sum_{y\in Y_1} f_1(y) f_2(y) P_{[1]}(y).$$

It follows easily that the set of functions $S = \{1, r_1, \ldots, r_m; r_1r_2, \ldots, r_{m-1}r_m; \ldots, r_1r_2\cdots r_m\}$ on Y_1 is **orthonormal** and, thus, is a basis in $V_\times$Since $g(y)$ is a function on Y_1, there exists a unique representation as a linear combination of functions in S, namely,

$$g(y) = \sum_{f\in S} < g, f > f.$$

$$\because < g, f >= \sum_{y\in Y_1} g(y) f(y) P_{[1]}(y) = \sum_{y\in Y_1} f(y) P(Y = y) = E_P[f]\,\forall f, \text{ and}$$

$$E_P[1] = 1, E_P\left[r_j\right] = 0, E_P\left[r_j r_k\right] = \rho_{jk}, \cdots, \text{ and } E_P\left[r_1\cdots r_m\right] = \rho_{12\cdots m}.$$

$$\therefore g(y) = \left(1 + \sum_{j<k} \rho_{jk} r_j r_k + \sum_{j<k<l} \rho_{jkl} r_j r_k r_l + \cdots + \rho_{12\cdots m} r_1 r_2 \cdots r_m\right).$$