BASIC PHD WRITTEN EXAMINATION IN BIOSTATISTICS

THEORY, SECTION 1

(9:00 AM–1:00 PM, July 25, 2016)

INSTRUCTIONS:

(a) This is a **CLOSED-BOOK** examination.

(b) The time limit for this examination is four hours.

(c) Answer any TWO (2) (BUT ONLY TWO) of the THREE (3) questions that follow.

(d) Put the answers to different questions on separate sets of paper.

(e) Put your exam code, **NOT YOUR NAME**, on each page. The same code will be used for Section 1 and Section 2 of the PhD Theory Exam. Please keep the code confidential and do not share this information with any students or faculty. Sharing your code with either students or faculty is viewed as a violation of the UNC honor code.

(f) Return the examination with a signed statement of the UNC honor pledge, separately from your answers. The pledge statement is given on the last page of the exam handout.

(g) In the questions to follow, you are required to answer only what is asked, and not to tell all you know about the topics involved.

1. (25 points) Let $X_1, ..., X_n$ be i.i.d from the following distribution

$$\begin{cases} 0 & \text{with probability } p, \\ \text{Uniform}[0, \theta] & \text{with probability } 1 - p. \end{cases}$$

First, we assume that p is a known constant in $(0, 1)$ and that $\theta > 0$ is the only parameter of interest.

(a) (5 points) Based on only one observation X_1, find all the unbiased estimators for θ and calculate their variances. Does UMVUE exist for θ? Justify your answer.

(b) (3 points) Based on n observations $X_1, ..., X_n$, let $X_{(n)} = \max\{X_1, ..., X_n\}$ be the maximal observation. Show that $(X_{(n)}, \sum_{i=1}^n I(X_i > 0))$ is a sufficient statistic for θ. Furthermore, show that $\widehat{\theta} = X_{(n)}$ maximizes the observed likelihood function.

(c) (5 points) What is the exact distribution of $\widehat{\theta}$? Compute $E[\widehat{\theta}]$ and $Var(\widehat{\theta})$ and show that $\widehat{\theta}$ is consistent for θ.

(d) (6 points) Derive the asymptotic distribution of $n(\widehat{\theta} - \theta)$.

Now assume that both p and θ are unknown.

(e) (6 points) Calculate the maximum likelihood estimator for p to obtain the maximum likelihood estimator for $E[X_1]$. Derive the asymptotic distribution for the latter after proper normalization.

2. (25 points) Suppose that $y_1, \ldots, y_n$ are independent binary random variables, where

$$P(y_i = 1|\beta_0, \beta_1, x_i) = \frac{\exp(\beta_0 + \beta_1 x_i)}{1 + \exp(\beta_0 + \beta_1 x_i)},$$

where $x_1, \ldots, x_n$ are fixed covariates and they are not all equal.

(a) (6 points) Suppose that (β_0, β_1) are both unknown and suppose we wish to test

$$H_0 : \beta_1 = 0 \text{ versus } H_1 : \beta_1 \neq 0.$$

Derive the Uniformly Most Powerful Unbiased (UMPU) α level test for this hypothesis and express the rejection region and critical value in the simplest possible form. Please note that there need not be a closed form for the distribution of the test statistic.

(b) (5 points) Using the UMPU conditional test from part (a), compute an explicit closed form for its conditional mean and conditional variance under the null hypothesis to find an explicit form for an asymptotically correct approximation to the UMPU test. You are allowed to assume that the conditional test statistic is asymptotically normal.

(c) (4 points) Derive the score test for the hypothesis in part (a), and compare its form to the approximate UMPU test derived in part (b).

(d) (6 points) Now consider the more general problem in which we have p covariates, and

$$P(y_i = 1|\boldsymbol{\beta}, \boldsymbol{x}_i) = \frac{\exp(\boldsymbol{x}_i'\boldsymbol{\beta})}{1 + \exp(\boldsymbol{x}_i'\boldsymbol{\beta})},$$

where $\boldsymbol{x}_i = (x_{i1}, \ldots, x_{ip})'$ is a $p \times 1$ vector of covariates, and $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)'$ is a $p \times 1$ vector of regression coefficients. Suppose we wish to test

$$H_0 : \ell'\boldsymbol{\beta} = \theta_0 \text{ versus } H_1 : \ell'\boldsymbol{\beta} \neq \theta_0,$$

where θ_0 is a specified scalar and ℓ is a specified and non-zero $p \times 1$ vector. Derive the UMPU size α test for this hypothesis and express the rejection region and critical value in the simplest possible form.

(e) (4 points) Describe in detail a non-parametric bootstrap algorithm for computing the exact p-value based on the UMPU test of part (d).

3. (25 points) Suppose $S \sim \text{Binomial}(n, p)$ and conditional on $S = s$, let $X_1, \dots, X_{s+1}$ be iid from a $N(\mu, 1)$ distribution. The value of n is known whereas (p, μ) are both unknown, $0 < p < 1$, and $-\infty < \mu < \infty$. We observe $(S, X_1, \dots, X_{S+1})$ and we wish to test $H_0 : \mu \leq 0$ versus $H_1 : \mu > 0$ at level α.

(a) (5 points) Write out the joint density of $(S, X_1, \dots, X_{S+1})$ and show that it belongs to a full-rank exponential family, and find the two dimensional complete sufficient statistic. Do the same thing for the special case that $\mu = 0$.

(b) (3 points) Derive the joint MLE's of (p, μ), denoted by $(\hat{p}, \hat{\mu})$.

(c) (5 points) Assuming that standard MLE theory applies, derive the joint asymptotic distribution of $(\hat{p}, \hat{\mu})$, properly normalized.

(d) (6 points) Let $\phi(S, X_1, \dots, X_{S+1})$ be *any* unbiased level α test of H_0 versus H_1. Write out what unbiasedness means for the power function $\beta(p, \mu)$ of such a test, and explain in detail why unbiasedness implies that $\beta(p, 0) = \alpha$ for all p.

(e) (6 points) Find the complete form of the UMPU test of H_0 versus H_1, including specification of the rejection region in terms of the sample mean of the X_i's and the $1 - \alpha$ quantile of a well known distribution.

2016 PhD Theory Exam, Section 1

Statement of the UNC honor pledge:

"In recognition of and in the spirit of the honor code, I certify that I have neither given nor received aid on this examination and that I will report all Honor Code violations observed by me."

(Signed) ______________________________________
NAME

(Printed) ______________________________________
NAME

1.

$X_1, \dots, X_n \sim p \cdot \delta_{\{0\}} + (1-p) \cdot unif[0, \theta]$

(a) Assume $T(X_1)$: UE for θ

Adopt a latent variable $Z_1 \sim Ber(p)$ s.t $\left(\begin{array}{l} Z_1 = 1 \Leftrightarrow X_1 = 0 \\ Z_1 = 0 \Leftrightarrow X_1 \sim unif[0,\theta] \end{array} \right)$

$$\mathbb{E}\, T(X_1) = \mathbb{E}\left[\mathbb{E}[T(X_1) \mid Z_1]\right]$$

$$= \mathbb{E}[T(X_1) \mid Z_1 = 1] \cdot P(Z_1 = 1) + \mathbb{E}[T(X_1) \mid Z_1 = 0] \cdot P(Z_1 = 0)$$

$X_1 \equiv 0$ $\qquad$ $X_1 \sim unif(0,\theta)$

$$= T(0) \cdot p + \int_0^\theta T(x) \frac{1}{\theta}\, dx \cdot (1-p)$$

Now, to make $T(X_1)$: UE for θ,

$\mathbb{E}\, T(X_1) = \theta$, $\forall \theta > 0$

$\Leftrightarrow T(0) \cdot p + \frac{1}{\theta} \int_0^\theta T(x) dx \cdot (1-p) = \theta$, $\forall \theta > 0$ (from the above)

$\Leftrightarrow \theta \cdot T(0) \cdot p + \int_0^\theta T(x) dx \cdot (1-p) = \theta^2$, $\forall \theta > 0$ (multiplying both sides by θ)

$\Rightarrow T(0) \cdot p + T(\theta) \cdot (1-p) = 2\theta$, $\forall \theta > 0$ (differentiating both sides by θ)

$\Rightarrow T(\theta) = \frac{1}{1-p}\left(2\theta - T(0) \cdot p\right)$, $\forall \theta > 0$

$\Rightarrow T(x) = \frac{1}{1-p}\left(2x - T(0) \cdot p\right)$, $x \in (0, \theta]$

$\therefore$ Any unbiased estimator $T(X_1)$ for θ should have the form of

$$T(x) = \begin{cases} \frac{1}{1-p}(2x - T(0) \cdot p), & x \in (0, \theta] \\ T(0) \ (\in \mathbb{R}), & x = 0 \end{cases}$$

Note that any UE for θ is fully specified the value at 0 $(= T(0))$.

(a) cont'd

Now, compute variance of $T(x_1)$.

$$\mathrm{Var}(T(X_1)) = \overset{①}{\mathbb{E}}\left[\mathrm{Var}(T(X_1)\mid Z_1)\right] + \overset{②}{\mathrm{Var}}\left(\mathbb{E}[T(X_1)\mid Z_1]\right)$$

①
$$\mathbb{E}\left[\mathrm{Var}(T(X_1)\mid Z_1)\right] = \underbrace{\mathrm{Var}(T(X_1)\mid Z_1=1)}_{X_1\equiv 0}\cdot P(Z_1=1) + \underbrace{\mathrm{Var}(T(X_1)\mid Z_1=0)}_{X_1\sim \mathrm{unif}[0,\theta]}\cdot P(Z_1=0)$$

$$= 0 \cdot p + \mathrm{Var}\left(\frac{1}{1-p}(2X_1 - T(0)\cdot p)\,\Big|\, Z_1=0\right)\cdot(1-p)$$

$X_1\sim \mathrm{unif}[0,\theta]$

$$= 0 + \left(\frac{2}{1-p}\right)^2 \mathrm{Var}(X_1\mid Z_1=0)\cdot(1-p)$$

$X_1\sim \mathrm{unif}[0,\theta]$

$$= \left(\frac{2}{1-p}\right)^2 \left\{\int_0^\theta x^2\cdot\frac{1}{\theta}\,dx - \left(\int_0^\theta x\cdot\frac{1}{\theta}\,dx\right)^2\right\}\cdot(1-p)$$

$$= \frac{4}{(1-p)}\cdot\frac{1}{12}\theta^2$$

$$= \frac{1}{3(1-p)}\theta^2$$

②
$$\mathbb{E}[T(X_1)\mid Z_1] = \begin{cases} \mathbb{E}[T(X_1)\mid Z_1=1], & Z_1=1 \ (X_1\equiv 0) \\ \mathbb{E}[T(X_1)\mid Z_1=0], & Z_1=0 \ (X_1\sim\mathrm{unif}[0,\theta]) \end{cases}$$

$$= \begin{cases} T(0), & Z_1=1 \quad (\because)\ X_1\equiv 0 \Rightarrow T(X_1)\equiv T(0) \\ \frac{2}{1-p}\mathbb{E}[X_1\mid Z_1=0] - \frac{p}{1-p}\cdot T(0), & Z_1=0 \quad (\because)\ T(X_1) = \frac{1}{1-p}(2X_1 - T(0)\cdot p) \end{cases}$$

$$= \begin{cases} T(0), & Z_1=1 \\ \frac{2}{1-p}\cdot\frac{\theta}{2} - \frac{p}{1-p}\cdot T(0), & Z_1=0 \quad (\because)\ Z_1=0 \Rightarrow X_1\sim\mathrm{unif}[0,\theta] \end{cases}$$

$$= T(0)\cdot Z_1 + \frac{1}{1-p}(\theta - p\cdot T(0))\cdot(1-Z_1) \quad (\text{using } Z_1 \text{ as an indicator})$$

$$= \frac{1}{1-p}(T(0)-\theta)\cdot Z_1 + \frac{1}{1-p}(\theta - p\cdot T(0))$$

$$\Rightarrow \mathrm{Var}\left(\mathbb{E}[T(X_1)\mid Z_1]\right) = \mathrm{Var}\left(\frac{1}{1-p}(T(0)-\theta)\cdot Z_1 + \frac{1}{1-p}(\theta - p\cdot T(0))\right)$$

$$= \left(\frac{T(0)-\theta}{1-p}\right)^2 \mathrm{Var}(Z_1)$$

$$= \frac{p}{1-p}\cdot(T(0)-\theta)^2 \quad (\because Z_1\sim \mathrm{Ber}(p))$$

(a) Cont'd

Finally, $\mathrm{Var}(T(X_1)) = \frac{1}{3(1-p)}\theta^2 + \frac{p}{1-p}(T(0)-\theta)^2$ //

Now, to check whether UMVUE for θ exists or not,

assume $T^*(X_1)$ with $T^*(0)=\alpha$ is a UMVUE for θ.

Then, $\mathrm{Var}_\theta(T^*(X_1)) \leq \mathrm{Var}_\theta(T(X_1))$, $\forall \theta>0$, $\forall T$: UE for θ.

Now, fix $\theta=\theta_0$.

Then, $\mathrm{var}_{\theta_0}(T^*(X_1)) = \frac{1}{3(1-p)}\theta_0^2 + \frac{p}{1-p}(\alpha-\theta_0)^2$

$\leq \mathrm{var}_{\theta_0}(T(X_1)) = \frac{1}{3(1-p)}\theta_0^2 + \frac{p}{1-p}(T(0)-\theta_0)^2$, for $\forall T(0)\in\mathbb{R}$.

(Note that "$T(0)$" fully specifies the UE $T(X_1)$)

For this, α should be θ_0.

It cannot be possible because α should vary by θ on which we want to achieve the minimal variance.

$\therefore$ NO UMVUE exists for θ //

(b) $f_{X_1}(x_1) = p^{I(x_1=0)} \cdot \left(\frac{1}{\theta}(1-p)\, I(x_1 \le \theta) \right)^{I(x_1>0)}$

Justification of $f_{X_1}(x_1)$ as a density (= Radon-Nykodym Derivative) of the given dist^n wrt dominating measure ν

$\nu = \mu + \delta_{\{0\}}$ (μ: Lebesgue, $\delta_{\{0\}}$: point measure at 0), $\delta_{\{0\}}(A) = I(0 \in A)$

$$\int_{[0,\theta]} f_{X_1}(x_1)\, d\nu(x_1) = \int_{[0,\theta]} f_{X_1}(x_1)\, d\mu(x_1) + \int_{[0,\theta]} f_{X_1}(x_1)\, d\delta_{\{0\}}(x_1)$$

$$= \int_{(0,\theta]} f_{X_1}(x_1)\, d\mu(x_1) + f_{X_1}(0) \cdot 1$$

($\because$) Lebesgue

$$= \frac{1}{\theta}(1-p)\cdot\theta + p = 1.$$

Likelihood

$$L(\theta) = \prod_i f_{X_1}(x_i) = p^{\sum_i I(x_i=0)} \cdot \left(\frac{1}{\theta}\right)^{\sum_i I(x_i>0)} \cdot (1-p)^{\sum_i I(x_i>0)} \cdot \prod_i \underline{I(x_i \le \theta)^{I(x_i>0)}} = I(0 \le x_i \le \theta)$$

	value
$x_1 = 0 \Rightarrow$	1
$0 < x_1 \le \theta \Rightarrow$	1
$x_1 > \theta \Rightarrow$	0

$$= p^{\sum_i I(x_i=0)} \cdot (1-p)^{\sum_i I(x_i>0)} \cdot \theta^{-\sum_i I(x_i>0)} \cdot I(0 \le x_{(1)}) \cdot I(x_{(n)} \le \theta)$$

$\therefore \left(X_{(n)}, \sum_{i=1}^n I(X_i>0)\right)$: SS for θ by Factorization Thm.

Likelihood $L(\theta) = p^{\sum_i I(x_i=0)} \cdot (1-p)^{\sum_i I(x_i>0)} \cdot \theta^{-\sum_i I(x_i>0)} \cdot I(0 \le x_{(1)}) \cdot I(x_{(n)} \le \theta)$

$\theta < x_{(n)} \Rightarrow L(\theta) = 0.$

$\theta \ge x_{(n)} \Rightarrow L(\theta) = C \cdot \left(\frac{1}{\theta}\right)^{\sum I(x_i > 0)}$: descreasing in θ.

$\therefore \hat{\theta} = X_{(n)}$ is MLE.

(c) For $0 < k \le \theta$,

$$P(X_{(n)} \le k) = \prod_{i=1}^{n} P(X_i \le k) = \prod_{i=1}^{n} \left(p + \frac{k}{\theta}(1-p)\right) = \left(p + \frac{k}{\theta}(1-p)\right)^n$$

$$(\because)\ P(X_i \le k) = P(X_i = 0) + P(0 < X_i \le k)$$
$$= p + (1-p) \cdot \int_0^k \frac{1}{\theta}\,dx$$

$$\Rightarrow f_{X_{(n)}}(k) = \frac{d}{dk} P(X_{(n)} \le k) = n \cdot \left(p + \frac{k}{\theta}(1-p)\right)^{n-1} \cdot \frac{1-p}{\theta} I(0 < k \le \theta) \quad : \text{density}$$

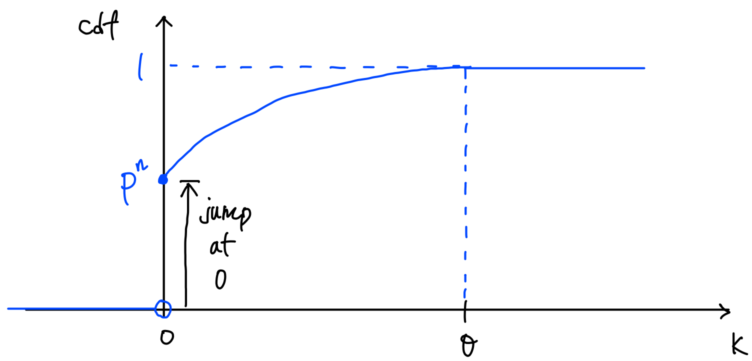
For $k = 0$,

$$P(X_{(n)} \le 0) = P(X_{(n)} = 0) = P(\forall X_i = 0) = p^n$$

For $k < 0$,

$$P(X_{(n)} \le k) = 0$$

$$\Rightarrow f_{X_{(n)}}(0) = p^n \quad : \text{point mass at } 0.$$



Now,

$$\mathbb{E}\hat{\theta} = \mathbb{E}X_{(n)} = 0 \cdot f_{X_{(n)}}(0) + \int_{(0,\theta]} k \cdot f_{X_{(n)}}(k)\,dk$$
$$= \int_{(0,\theta]} k \cdot \frac{n}{\theta}(1-p) \cdot \left(p + \frac{k}{\theta}(1-p)\right)^{n-1} dk$$
$$= \int_p^1 \frac{\theta}{1-p}(y-p) \cdot \frac{n(1-p)}{\theta} \cdot y^{n-1} \cdot \frac{\theta}{1-p}\,dy \quad \left(\text{Letting } y := p + \frac{k}{\theta}(1-p)\right)$$
$$= \frac{n\theta}{1-p} \int_p^1 y^n - p y^{n-1}\,dy$$
$$= \frac{n\theta}{1-p} \left[\frac{1}{n+1} y^{n+1} - \frac{p}{n} y^n\right]_p^1$$
$$= \frac{n\theta}{1-p} \left(\frac{1}{n+1} - \frac{p}{n} - \frac{p^{n+1}}{n+1} + \frac{p^{n+1}}{n}\right)$$
$$= \theta - \frac{\theta}{1-p} \cdot \frac{1-p^{n+1}}{n+1}$$

(c) cont'd

$$Var(\hat\theta) = Var(X_{(n)}) = \mathbb{E}X_{(n)}^2 - (\mathbb{E}X_{(n)})^2$$

$$\mathbb{E}X_{(n)}^2 = \int_0^\theta k^2 \frac{n(1-p)}{\theta}\left(p+\frac{k}{\theta}(1-p)\right)^{n-1} dk$$

$$= \int_p^1 \left(\frac{\theta}{1-p}(y-p)\right)^2 \frac{n(1-p)}{\theta} y^{n-1} \frac{\theta}{1-p} dy$$

$$= n\left(\frac{\theta}{1-p}\right)^2 \int_p^1 y^{n+1} - 2py^n + p^2 y^{n-1} dy$$

$$= n\left(\frac{\theta}{1-p}\right)^2 \left[\frac{1}{n+2}y^{n+2} - \frac{2p}{n+1}y^{n+1} + \frac{p^2}{n}y^n\right]_p^1$$

$$= n\left(\frac{\theta}{1-p}\right)^2 \left[\frac{1}{n+2} - \frac{2p}{n+1} + \frac{p^2}{n} - \frac{p^{n+2}}{n+2} + \frac{2p^{n+2}}{n+1} - \frac{p^{n+2}}{n}\right]$$

$$Var(\hat\theta) = n\left(\frac{\theta}{1-p}\right)^2 \left[\frac{1}{n+2} - \frac{2p}{n+1} + \frac{p^2}{n} - \frac{p^{n+2}}{n+2} + \frac{2p^{n+2}}{n+1} - \frac{p^{n+2}}{n}\right] - \left(n\frac{\theta}{1-p}\left(\frac{1}{n+1} - \frac{p}{n} - \frac{p^{n+1}}{n+1} + \frac{p^{n+1}}{n}\right)\right)^2$$

$$= n\left(\frac{\theta}{1-p}\right)^2 \left\{ \left[\quad\right] - n\left(\frac{1}{n+1} - \frac{p}{n} - \frac{p^{n+1}}{n+1} + \frac{p^{n+1}}{n}\right)^2 \right\}$$

$$= \quad 〃 \quad \left\{ \left[\quad\right] - n\left(\frac{1-p}{n} + \frac{1-p^{n+1}}{n(n+1)}\right)^2 \right\}$$

To show consistency of $\hat\theta$,

Markov : $$P(|\hat\theta-\theta| \geq \epsilon) \leq \frac{\mathbb{E}|\hat\theta-\theta|^2}{\epsilon^2} = \frac{\mathbb{E}\hat\theta^2 - 2\theta\mathbb{E}\hat\theta + \theta^2}{\epsilon^2}$$

$$= \frac{1}{\epsilon^2}\left\{ n\left(\frac{\theta}{1-p}\right)^2 \left(\frac{1}{n+2} - \frac{2p}{n+1} + \frac{p^2}{n} - \frac{p^{n+2}}{n+2} + \frac{2p^{n+2}}{n+1} - \frac{p^{n+2}}{n}\right) - 2\theta\cdot\left(\theta - \frac{\theta}{1-p}\frac{1-p^{n+1}}{n+1}\right) + \theta^2 \right\}$$

$$\xrightarrow[n\to\infty]{} \frac{1}{\epsilon^2}\left\{\left(\frac{\theta}{1-p}\right)^2(1-2p+p^2) - 2\theta^2 + \theta^2\right\} = 0$$

$\therefore$ $\hat\theta \to_p \theta$ proved.

sol 1)

Markov : $$P(|\hat\theta - \mathbb{E}\hat\theta| > \epsilon) \leq \frac{Var(\hat\theta)}{\epsilon^2} \downarrow 0$$

$$\mathbb{E}\hat\theta = \frac{n\theta}{1-p}\left(\frac{1}{n+1} - \frac{p}{n} - \frac{p^{n+1}}{n+1} + \frac{p^{n+1}}{n}\right)$$

$$= \frac{n\theta}{1-p}\left(\frac{1-p}{n} + \frac{1}{n+1} - \frac{1}{n} - \frac{p^{n+1}}{n+1} + \frac{p^{n+1}}{n}\right)$$

$$= \theta + \frac{\cancel{n}\theta}{1-p}\cdot\frac{1}{\cancel{n}(n+1)}\left(-1+p^{n+1}\right) \xrightarrow[n]{} \theta$$

<WTS>

$$P(|\hat\theta-\theta| > \epsilon) \to 0$$

$$|\hat\theta-\theta| \leq |\hat\theta - \mathbb{E}\hat\theta| + \underline{|\mathbb{E}\hat\theta - \theta|} \to 0$$

$\exists N$ s.t. $|\mathbb{E}\hat\theta - \theta| < \frac{\epsilon}{2}$, $\forall n > N$

$$\to P(|\hat\theta-\theta| > \epsilon) \leq P(|\hat\theta - \mathbb{E}\hat\theta| + |\mathbb{E}\hat\theta - \theta| > \epsilon)$$

$$\leq P(|\hat\theta - \mathbb{E}\hat\theta| + \frac{\epsilon}{2} > \epsilon)$$

$$= P(|\hat\theta - \mathbb{E}\hat\theta| > \frac{\epsilon}{2}) \to 0.$$

sol 2)

$$P(|\hat\theta-\theta| > \epsilon) \leq \frac{\mathbb{E}((\hat\theta-\theta)^2)}{\epsilon^2} = \frac{\mathbb{E}\hat\theta^2 - 2\theta\mathbb{E}\hat\theta + \theta^2}{\epsilon^2} \xrightarrow[n]{} 0$$

(d)

$Y_n = n(\hat\theta - \theta)$ $\quad \hat\theta \in [0,\theta] \Rightarrow Y_n \in [-n\theta, 0]$

$$P(Y_n \le k) = P(n(\hat\theta-\theta) \le k) = P\left(\hat\theta \le \theta + \frac{k}{n}\right) = \left(p + \frac{\theta + \frac{k}{n}}{\theta}(1-p)\right)^n = \left(1 + \frac{k}{n\theta}(1-p)\right)^n = \left(1 + \frac{\frac{k}{\theta}(1-p)}{n}\right)^n$$

$k \in [-n\theta, 0]$

$n \to \infty$: $P(Y_n \le k) \to e^{\frac{1-p}{\theta}k}$, $k \le 0$

$Z_n := -Y_n \sim \exp\left(\frac{1-p}{\theta}\right)$

$$P(Z_n \le t) = P(Y_n \ge -t) = 1 - P(Y_n \le -t) \to \underset{\text{"}t>0\text{"}}{1 - e^{-\frac{1-p}{\theta}t}} \quad : \text{cdf of } Exp\left(\frac{1-p}{\theta}\right)$$

(e)

Likelihood $L(\theta) = p^{\sum_i I(x_i=0)} \cdot (1-p)^{\sum_i I(x_i>0)} \cdot \theta^{-\sum_i I(x_i>0)} \cdot I(0 \le x_{(1)}) \cdot I(x_{(n)} \le \theta)$

$$\Rightarrow \frac{\partial L}{\partial p} \overset{set}{=} 0 \Rightarrow \hat p = \frac{\sum I(x_i=0)}{\sum I(x_i=0) + \sum I(x_i>0)} = \frac{\sum I(x_i=0)}{n}$$

Parameter of Interest : $EX_1 = \frac{\theta}{2}(1-p)$

$\Rightarrow \widehat{EX_1} = \frac{\hat\theta}{2}(1-\hat p)$ $(\because)$ Invariance of MLE wrt 1-1 function

Note that $I(X_i=0) \overset{iid}{\sim} Ber(p)$

CLT : $\sqrt{n}\left(\frac{1}{n}\sum_i I(X_i=0) - E\,I(X_i=0)\right) \to_d N(0, Var(I(X_i=0)))$

$\Rightarrow$ $\underline{\underline{\sqrt{n}(\hat p - p) \to_d N(0, p(1-p))}}$ $(*)$

$$\therefore \sqrt{n}\left(\hat\theta\frac{1-\hat p}{2} - \theta\frac{1-p}{2}\right) = n(\hat\theta - \theta)\frac{1-\hat p}{2\sqrt{n}} + \sqrt{n}\,\theta\left(\frac{1-\hat p}{2} - \frac{1-p}{2}\right) \quad \left(\text{substract \& add } \sqrt{n}\,\theta\frac{1-\hat p}{2}\right)$$

$$\to_d -T \cdot 0 - \frac{\theta}{2}\underbrace{\underline{\underline{\sqrt{n}(\hat p - p)}}}_{(*)} \quad (\because)\ \hat p \to_p p,\ \frac{1}{\sqrt{n}} \to 0 \Rightarrow \frac{1-\hat p}{2\sqrt{n}} \to_p 0$$

$$\to_d N\left(0, \frac{\theta^2}{4}p(1-p)\right)$$

<The reason why we can't use usual asymp. normality theory>

$$f_{X_1}(x_1) = p^{I(x_1=0)}\left(\frac{1}{\theta}\right)^{I(x_1>0)}(1-p)^{I(x_1>0)}$$

$I(0 \le x_1 \le \theta)$ depends on "θ" $\Rightarrow$ NO common support.

2.

$$P(y_i=1 \mid \beta_0, \beta_1, x_i) = \frac{\exp(\beta_0+\beta_1 x_i)}{1+\exp(\beta_0+\beta_1 x_i)}$$

(a)

likelihood

$$L(\beta_0,\beta_1) = \prod_{i=1}^{n} P(y_i=1 \mid \beta_0,\beta_1,x_i)^{y_i}\left(1-P(y_i=1\mid \beta_0,\beta_1,x_i)\right)^{1-y_i}$$

$$= \prod_{i=1}^{n} \frac{\exp\left(y_i(\beta_0+\beta_1 x_i)\right)}{1+\exp(\beta_0+\beta_1 x_i)}$$

$$= \exp\left\{\sum_{i=1}^{n} y_i\cdot\beta_0 + \sum_{i=1}^{n} x_i y_i\cdot\beta_1 - \sum_{i=1}^{n}\log\left(1+\exp(\beta_0+\beta_1 x_i)\right)\right\} \quad : \text{exp'l family.}$$

$U(Y) = \sum_{i=1}^{n} x_i Y_i$, $T(Y) = \sum_{i=1}^{n} Y_i$: SS for β_1, β_0 respectively.

UMPU α level test for $H_0: \beta_1=0$ vs $H_1: \beta_1\neq 0$ is given by

(*)

$$\phi(Y) = \begin{cases} 1 & U<C_1(t) \text{ or } U>C_2(t) \\ \gamma_i & U=C_i(t),\ i=1,2 \\ 0 & \text{o.w} \end{cases} \quad (**)$$

$$E_{\beta_1=0}[\phi(U)\mid T=t] = \alpha$$

$$E_{\beta_1=0}[U\phi(U)\mid T=t] = \alpha\cdot E_{\beta_1=0}[U\mid T=t]$$

where t is an arbitrary value of T

$Y_i \sim Ber(p_i)$, $p_i = \exp(\beta_0+\beta_1 x_i)/\{1+\exp(\beta_0+\beta_1 x_i)\}$

At most 2^n values of $U(Y)$ are possible.

Under $H_0: \beta_1=0$, $Y_i \overset{iid}{\sim} Ber(p_0)$, $p_0 = \frac{\exp(\beta_0)}{1+\exp(\beta_0)}$. $\Rightarrow T \overset{H_0}{\sim} B(n,p_0)$

$$\therefore P_{\beta_1=0}(U=u\mid T=t) = \frac{P_{\beta_1=0}(U=u, T=t)}{P_{\beta_1=0}(T=t)} = \frac{\sum_{y\in A(u,t)}\prod_{i=1}^{n} p_0^{y_i}(1-p_0)^{1-y_i}}{\binom{n}{t}p_0^{t}(1-p_0)^{n-t}},$$

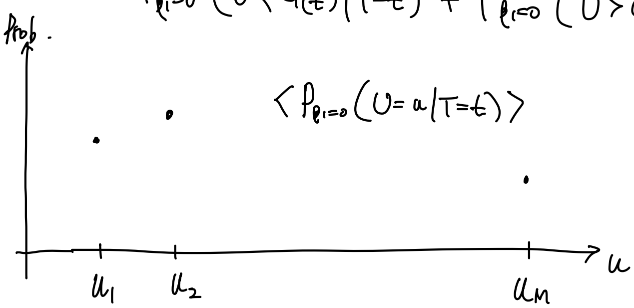
$A(u,t) = \{(y_1,\cdots,y_n) : \sum y_i = t, \sum x_i y_i = u\}$

$$= \frac{|A(u,t)|\cdot p_0^{t}(1-p_0)^{n-t}}{\binom{n}{t}p_0^{t}\cdot(1-p_0)^{n-t}} = \frac{|A(u,t)|}{\binom{n}{t}}$$

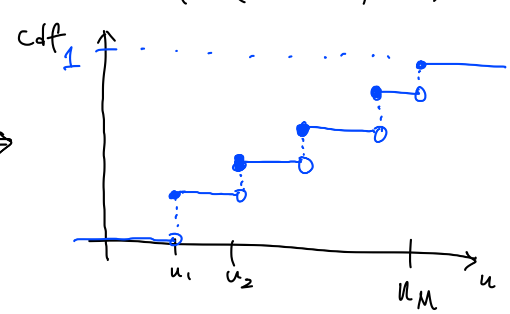
$|A(u,t)|$ ← # of elements of the set.

$$\alpha = E_{\beta_1=0}[\phi(U)\mid T=t]$$

$$= P_{\beta_1=0}(U<C_1(t)\mid T=t) + P_{\beta_1=0}(U>C_2(t)\mid T=t) + \gamma_1\cdot P_{\beta_1=0}(U=C_1(t)\mid T=t) + \gamma_2\cdot P_{\beta_1=0}(U=C_2(t)\mid T=t)$$



$\Rightarrow$



$u_1,\cdots,u_M$: possible value of u given $\sum y_i = t$.

→ Let $\mathcal{U}$

$$\alpha\cdot E_{\beta_1=0}[U\mid T=t] = \alpha\cdot\sum_{u\in\mathcal{U}} u\cdot\frac{|A(u,t)|}{\binom{n}{t}}$$

$$= E_{\beta_1=0}[U\phi(U)\mid T=t] = \sum_{u<C_1(t)} u\cdot\frac{|A(u,t)|}{\binom{n}{t}} + \sum_{u>C_2(t)} u\cdot\frac{|A(u,t)|}{\binom{n}{t}} + \gamma_1\cdot\frac{|A(C_1(t),t)|}{\binom{n}{t}}\cdot I(C_1(t)\in\mathcal{U}) + \gamma_2\cdot\frac{|A(C_2(t),t)|}{\binom{n}{t}}\, I(C_2(t)\in\mathcal{U})$$

We can find such $C_1(t)$ and $C_2(t)$ among $u_1,\cdots,u_M$.

(b)

$$\mathbb{E}_{H_0}[U \mid T=t] = \mathbb{E}_{H_0}\left[\sum_i x_i Y_i \,\Big|\, \sum_i Y_i = t\right] = \sum_i x_i \mathbb{E}_{H_0}\left[Y_i \,\Big|\, \sum_i Y_i = t\right]$$

$$= \left(\sum_i x_i\right) \cdot \mathbb{E}_{H_0}\left[Y_1 \,\Big|\, \sum_i Y_i = t\right] \quad \left(\text{symmetricity}: \mathbb{E}_{H_0}\left[Y_i \,\Big|\, \sum_i Y_i = t\right] \text{ all same.}\right)$$

$$= \left(\sum_i x_i\right) \cdot \frac{1}{n} \mathbb{E}_{H_0}\left[\sum_i Y_i \,\Big|\, \sum_i Y_i = t\right] \quad (\text{once again, symmetricity})$$

$$= \sum_i x_i \cdot \frac{t}{n}$$

$$\mathrm{Var}_{H_0}(U \mid T=t) = \mathbb{E}_{H_0}\left[\left(\sum_i x_i Y_i\right)^2 \,\Big|\, \sum_i Y_i = t\right] - \left(\sum_i x_i \cdot \frac{t}{n}\right)^2$$

$$= \sum_i x_i^2 \mathbb{E}_{H_0}\left[Y_i^2 \,\Big|\, \sum_i Y_i = t\right] + \sum_{i \neq j} x_i x_j \mathbb{E}_{H_0}\left[Y_i Y_j \,\Big|\, \sum_i Y_i = t\right] - \left(\sum_i x_i \frac{t}{n}\right)^2$$

$$= \left(\sum_i x_i^2\right) \cdot \mathbb{E}_{H_0}\left[Y_1^2 \,\Big|\, \sum_i Y_i = t\right] + \left(\sum_{i \neq j} x_i x_j\right) \cdot \mathbb{E}_{H_0}\left[Y_1 Y_2 \,\Big|\, \sum_i Y_i = t\right] - \text{''} \quad (\text{symmetricity})$$

Here, note that $Y_i \overset{H_0}{\sim}$ iid $\mathrm{Ber}(p_0)$ in (a).

Then,

$$P\left(Y_1 = 1 \,\Big|\, \sum_i Y_i = t\right) = \frac{P\left(Y_1 = 1, \sum_{i=2}^n Y_i = t-1\right)}{P\left(\sum_{i=1}^n Y_i = t\right)} = \frac{p_0 \cdot \binom{n-1}{t-1} p_0^{t-1} (1-p_0)^{n-t}}{\binom{n}{t} p_0^t (1-p_0)^{n-t}} \quad (\because) \sum_{i=2}^n Y_i \sim B(n-1, p_0)$$

$$= \frac{t}{n}$$

$$P\left(Y_1 = 0 \,\Big|\, \sum_i Y_i = t\right) = 1 - P\left(Y_1 = 1 \,\Big|\, \sum_i Y_i = t\right) = \frac{n-t}{n}$$

$$\Rightarrow \mathbb{E}_{H_0}\left[Y_1^2 \,\Big|\, \sum_i Y_i = t\right] = 1^2 \cdot P\left(Y_1 = 1 \,\Big|\, \sum_i Y_i = t\right) = \frac{t}{n}$$

$$\mathbb{E}_{H_0}\left[Y_1 Y_2 \,\Big|\, \sum_i Y_i = t\right] = 1 \cdot 1 \cdot P\left(Y_1 = 1, Y_2 = 1 \,\Big|\, \sum_i Y_i = t\right) = \frac{P\left(Y_1 = 1, Y_2 = 1, \sum_{i=3}^n Y_i = t-2\right)}{P\left(\sum_i Y_i = t\right)}$$

$$= \frac{p_0 \cdot p_0 \cdot \binom{n-2}{t-2} p_0^{t-2} \cdot (1-p_0)^{n-t}}{\binom{n}{t} p_0^t (1-p_0)^{n-t}} = \frac{t(t-1)}{n(n-1)}$$

$$\therefore \mathbb{E}_{H_0}[U \mid T=t] = \sum_i x_i \frac{t}{n}$$

$$\mathrm{Var}_{H_0}(U \mid T=t) = \left(\sum_i x_i^2\right) \cdot \frac{t}{n} + \overbrace{\left(\sum_{i \neq j} x_i x_j\right)}^{\left(\sum x_i\right)^2 - \sum x_i^2} \frac{t(t-1)}{n(n-1)} - \left(\sum x_i\right)^2 \frac{t^2}{n^2}$$

$$= \left(\sum_i x_i^2\right)\left(\frac{t}{n} - \frac{t(t-1)}{n(n-1)}\right) - \left(\sum x_i\right)^2 \cdot \left(\frac{t^2}{n^2} - \frac{t(t-1)}{n(n-1)}\right)$$

$$= \sum_i x_i^2 \frac{t}{n} \frac{n-t}{n-1} - \left(\sum x_i\right)^2 \frac{t}{n} \frac{n-t}{n(n-1)}$$

$$= \frac{t}{n} \frac{n-t}{n} \frac{n^2}{n-1} \left\{ \frac{1}{n} \sum x_i^2 - \left(\frac{1}{n} \sum x_i\right)^2 \right\}$$

$$= \frac{n^2}{n-1} \cdot \left(\frac{t}{n}\right)\left(1 - \frac{t}{n}\right) \left\{ \overline{x^2} - \bar{x}^2 \right\} \quad \left(= \frac{n^2}{n-1} \bar{Y}(1-\bar{Y}) \left\{ \overline{x^2} - \bar{x}^2 \right\} \right)$$

(b) Now, assuming $U|T=t \overset{H_0}{\sim} N(E,V)$, $\left(E=\mathbb{E}_{H_0}[U|T=t],\ V=\mathrm{Var}_{H_0}(U|T=t)\right)$

UMPU test in (a) $\xrightarrow{\text{Approximation}}$ Asymptotic Normal test

(i)
$$\phi(Y)=\begin{cases}1 & U<c_1(t) \text{ or } U>c_2(t)\\ \gamma_i & U=c_i(t)\\ 0 & \text{o.w}\end{cases}$$

(i)' $\phi(Y)=I\left(U<c_1(t) \text{ or } U>c_2(t)\right)$

($\because$) $U|T=t \sim$ Normal $\Rightarrow$ continuous

(ii) $\mathbb{E}_{\beta_1=0}[\phi(U)|T=t]=\alpha$

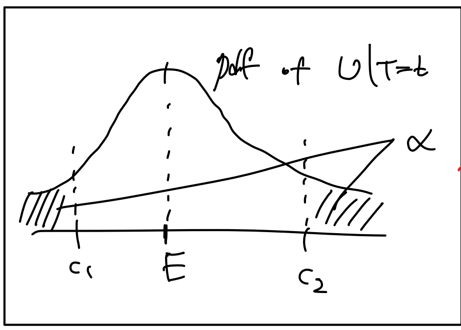
$$\mathbb{E}_{\beta_1=0}[U\phi(U)|T=t]=\alpha\cdot\mathbb{E}_{\beta_1=0}[U|T=t]$$

(ii)' $P_{\beta_1=0}\left(U<c_1(t) \text{ or } U>c_2(t)\right)=\alpha$

$$\mathbb{E}_{\beta_1=0}[UI(\quad " \quad)|T=t]=\alpha\cdot\mathbb{E}_{\beta_1=0}[U|T=t]$$

$\therefore$ (i)' & (ii)' : Find c_1 & c_2 s.t

$$P(U<c_1 \text{ or } U>c_2|T=t)=\alpha \ \ \& \ \ \mathbb{E}[UI(U<c_1 \text{ or } U>c_2)|T=t]=\alpha\cdot\mathbb{E}[U|T=t]$$

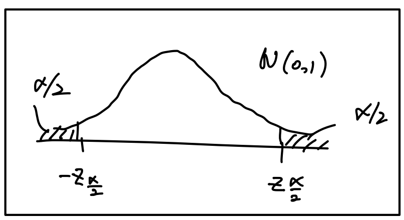


Assume c_1 & c_2 : symmetric on E

$1-P(c_1\le U\le c_2)=\alpha \Rightarrow P(c_1\le U\le c_2)=1-\alpha$

Approximation

$$\Rightarrow P\left(\frac{c_1-E}{\sqrt{V}}\le \underbrace{\frac{U-E}{\sqrt{V}}}_{N(0,1)}\le\frac{c_2-E}{\sqrt{V}}\right)=1-\alpha$$



$$\Rightarrow c_1=E-z_{\frac{\alpha}{2}}\cdot\sqrt{V}$$
$$c_2=E+z_{\frac{\alpha}{2}}\cdot\sqrt{V} \ //$$

Now, check whether these c_1 & c_2 achieves (ii)'

$$\mathbb{E}_{\beta_1=0}[UI(U<c_1 \text{ or } U>c_2)|T=t]$$
$$=\mathbb{E}_{\beta_1=0}[U|T=t]-\mathbb{E}_{\beta_1=0}[UI(c_1\le U\le c_2)|T=t]$$
$$=\quad E \quad -\mathbb{E}\left[(E+\sqrt{V}\cdot Z)I(-z_{\frac{\alpha}{2}}\le Z\le z_{\frac{\alpha}{2}})\right] \qquad \left(Z\overset{d}{=}\frac{U-E}{\sqrt{V}}\Big|T=t \sim N(0,1)\right)$$
$$=E-E\cdot\underbrace{P(-z_{\frac{\alpha}{2}}\le Z\le z_{\frac{\alpha}{2}})}_{=1-\alpha}-\underbrace{\sqrt{V}\cdot\mathbb{E}\left[ZI(-z_{\frac{\alpha}{2}}\le Z\le z_{\frac{\alpha}{2}})\right]}_{=0 \text{ by symmetricity}}$$
$$=\alpha\cdot E$$
$$=\alpha\cdot\mathbb{E}_{\beta_1=0}[U|T=t] \ //$$

$\therefore c_1=E-z_{\frac{\alpha}{2}}\cdot\sqrt{V},\ c_2=E+z_{\frac{\alpha}{2}}\cdot\sqrt{V}$ achieves (ii)' //

$\therefore$ Approximated test: $\phi(Y)=I\left(\left|\sum x_iY_i-\mathbb{E}_{H_0}[U|T=t]\right|/\mathrm{Var}_{H_0}(U|T=t)^{1/2}>z_{\frac{\alpha}{2}}\right)$

$$=I\left(\left|\sum x_iY_i-\bar{x}\cdot\sum Y_i\right|/\left(\frac{n^2}{n-1}\cdot\bar{Y}(1-\bar{Y})(\overline{x^2}-\bar{x}^2)\right)^{1/2}>z_{\frac{\alpha}{2}}\right)$$

(c)

Score test for $H_0: \beta_1 = 0 \Leftrightarrow H_0: L\beta = 0$, $L = [0 \ 1]$, $\beta = \begin{bmatrix}\beta_0 \\ \beta_1\end{bmatrix}$

Assume. $X_1 \sim X_n \sim iid\ P(\cdot\,;\theta)$

$\ell(\theta) = \frac{1}{n}\sum_{i=1}^{n} \log P(X_i;\theta)$ $\quad \ell_i(\theta) = \log P(X_i;\theta) \Rightarrow \ell(\theta) = \frac{1}{n}\sum_{i=1}^{n}\ell_i(\theta)$

$\dot{\ell}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial}{\partial\theta}\log P(X_i;\theta)$

$\ddot{\ell}(\theta) = \frac{1}{n}\sum_{i=1}^{n} \frac{\partial^2}{\partial\theta^2}\log P(X_i;\theta)$

Bartlett's identity: $\mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log P(X_1;\theta)\right] = 0$

$\mathrm{Var}_\theta[\quad " \quad] = \mathbb{E}_\theta\left[\left(\frac{\partial}{\partial\theta}\log P(X_1;\theta)\right)^2\right] = -\mathbb{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}\log P(X_1;\theta)\right] = I_1(\theta)$

CLT: $\sqrt{n}\left(\dot{\ell}(\theta) - \mathbb{E}_\theta\left[\frac{\partial}{\partial\theta}\log P(X_1;\theta)\right]\right) \rightarrow_d N\left(0, \mathrm{Var}_\theta\left[\frac{\partial}{\partial\theta}\log P(X_1;\theta)\right]\right)$

$\Rightarrow \sqrt{n}\,\dot{\ell}(\theta) \rightarrow_d N(0, I_1(\theta))$

WLLN: $\ddot{\ell}(\theta) = \frac{1}{n}\sum \frac{\partial^2}{\partial\theta^2}\log P(X_i;\theta) \rightarrow_p \mathbb{E}_\theta\left[\frac{\partial^2}{\partial\theta^2}\log P(X_1;\theta)\right] = -I_1(\theta)$

Taylor Exp: $\dot{\ell}(\theta) = \dot{\ell}(\hat{\theta}) + \ddot{\ell}(\hat{\theta})(\theta - \hat{\theta}) + \underbrace{\frac{1}{2}(\theta-\hat{\theta})^T \dddot{\ell}(\tilde{\theta})(\theta-\hat{\theta})}_{\text{Tensor product}}$

$\Rightarrow \dot{\ell}(\theta) = \left[\ddot{\ell}(\hat{\theta}) + \frac{1}{2}(\theta-\hat{\theta})^T\dddot{\ell}(\tilde{\theta})\right](\theta-\hat{\theta})$

$= \left[\ddot{\ell}(\theta) + o_p(1) + O_p(M\cdot\|\hat{\theta}-\theta\|_1)\right](\theta-\hat{\theta})$

$= \left[-I_1(\theta) + o_p(1) + o_p(1)\right](\theta-\hat{\theta})$ $\quad \rightarrow 0$ from consistency

$\Rightarrow \sqrt{n}(\hat{\theta}-\theta) = I_1(\theta)^{-1}\sqrt{n}\cdot\dot{\ell}(\theta) + o_p(1)$

$\rightarrow_d N(0, I_1(\theta)^{-1})$ //

<u>Walk's phenomenon</u>

① $(\sqrt{n}(\hat{\theta}-\theta))^T I_1(\theta)(\sqrt{n}(\hat{\theta}-\theta)) \rightarrow_d \chi^2_{\dim(\theta)}$

$\Rightarrow (\hat{\theta}-\theta)^T\cdot(n I_1(\theta))\cdot(\hat{\theta}-\theta) \rightarrow_d \chi^2_{\dim(\theta)}$

$\Rightarrow (\hat{\theta}-\theta)^T\left[-\sum_{i=1}^{n}\frac{\partial^2}{\partial\theta^2}\log P(X_i;\theta)\right](\hat{\theta}-\theta) \rightarrow_d \chi^2_{\dim(\theta)}$: <u>Wald</u>

or

$(\hat{\theta}-\theta)^T(n I_1(\hat{\theta}))(\hat{\theta}-\theta) \rightarrow_d \chi^2_{\dim(\theta)}$.

② $(\sqrt{n}\,\dot{\ell}(\theta))^T I_1(\theta)^{-1}(\sqrt{n}\,\dot{\ell}(\theta)) \rightarrow_d \chi^2_{\dim(\theta)}$

$\Rightarrow n\cdot\dot{\ell}(\theta)^T I_1(\theta)^{-1}\dot{\ell}(\theta) \rightarrow_d \chi^2_{\dim(\theta)}$

$\Rightarrow n\cdot\dot{\ell}(\hat{\theta})^T I_1(\hat{\theta})^{-1}\dot{\ell}(\hat{\theta}) \rightarrow_d \chi^2_{\dim(\theta)}$: <u>Score</u>

or

$n\cdot\dot{\ell}(\hat{\theta})^T\left[-\sum_i \frac{\partial^2}{\partial\theta^2}\log P(X_i;\theta)\big|_{\theta=\hat{\theta}}\right]^{-1}\dot{\ell}(\hat{\theta}) \rightarrow_d \chi^2_{\dim(\theta)}$

<u>Tests based on phenomena</u>

① Wald Test for $H_0: H(\theta) = 3$ vs H_1: NOT

$H(\theta) = H(\hat{\theta}) + \dot{H}(\hat{\theta})(\theta-\hat{\theta}) + \frac{1}{2}(\theta-\hat{\theta})^T\ddot{H}(\tilde{\theta})(\theta-\hat{\theta})$

$\simeq H(\hat{\theta}) + \dot{H}(\hat{\theta})(\theta-\hat{\theta})$

$\Rightarrow \hat{\theta}-\theta \simeq \dot{H}(\hat{\theta})^{-1}(H(\hat{\theta})-3)$

$\Rightarrow W_n := (H(\hat{\theta})-3)^T(\dot{H}(\hat{\theta})^{-1})^T(n I_1(\hat{\theta}))$

$\dot{H}(\hat{\theta})^{-1}\cdot(H(\hat{\theta})-3) \rightarrow_d \chi^2_{df}$

② Score Test for $H_0: H(\theta) = 3$ vs H_1: NOT

Find $\hat{\theta}^0$: MLE of θ under $H_0: H(\theta) = 3$.

Then,

$SC_n := n\cdot\dot{\ell}(\theta)^T\left[-\sum_i \frac{\partial^2}{\partial\theta}\log P(X_i;\theta)\right]^{-1}\dot{\ell}(\theta)\Big|_{\theta=\hat{\theta}^0}$

$\rightarrow_d \chi^2_{df}$.

$\boxed{\text{NOTE}}$ $df = \dim(\Theta) - \dim\{\theta\in\Theta : H(\theta) = 3\}$.

(c) cont'ed

$$l(\beta_0,\beta_1) = \sum_{i=1}^{n} \log p(y_i \mid \beta_0,\beta_1,x_i) = \sum_{i=1}^{n} y_i \cdot (\beta_0+\beta_1 x_i) - \log(1+\exp(\beta_0+\beta_1 x_i)) . \quad \theta := \begin{bmatrix} \beta_0 \\ \beta_1 \end{bmatrix}.$$

$$\frac{\partial l}{\partial \beta_0} = \sum_{i=1}^{n} y_i - \frac{\exp(\beta_0+\beta_1 x_i)}{1+\exp(\beta_0+\beta_1 x_i)} = \sum_i (y_i - p_i) . \qquad p_i := \frac{\exp(\beta_0+\beta_1 x_i)}{1+\exp(\beta_0+\beta_1 x_i)}$$

$$\frac{\partial l}{\partial \beta_1} = \sum_{i=1}^{n} y_i \cdot x_i - \quad '' \quad \cdot x_i = \sum_i (y_i - p_i)\cdot x_i .$$

$$\frac{\partial p_i}{\partial \beta_0} = \frac{\exp(\cdots)\cdot(1+\exp(\cdots)) - \exp(\cdots)\cdot\exp(\cdots)}{(1+\exp(\cdots))^2} = p_i(1-p_i)$$

$$\frac{\partial p_i}{\partial \beta_1} = \frac{\exp(\cdots)\cdot x_i\,(1+\exp(\cdots)) - \exp(\cdots)\cdot\exp(\cdots)\cdot x_i}{(\qquad)^2} = p_i(1-p_i)\,x_i$$

$$\frac{\partial^2 l}{\partial \beta_0^2} = -\sum_i \frac{\partial p_i}{\partial \beta_0} = -\sum_i p_i(1-p_i)$$

$$\frac{\partial^2 l}{\partial \beta_1 \partial \beta_0} = -\sum_i \frac{\partial p_i}{\partial \beta_1} = -\sum_i p_i(1-p_i)x_i$$

$$\frac{\partial^2 l}{\partial \beta_1^2} = -\sum_i \frac{\partial p_i}{\partial \beta_1}\cdot x_i = -\sum_i p_i(1-p_i)x_i^2$$

$$\therefore \dot{l}(\theta) = \begin{bmatrix} \frac{\partial l}{\partial \beta_0} \\ \frac{\partial l}{\partial \beta_1} \end{bmatrix} = \begin{bmatrix} \sum_i (y_i - p_i) \\ \sum_i (y_i - p_i)x_i \end{bmatrix}$$

$$I(\theta) = -\mathbb{E}_\theta\left[\frac{\partial^2}{\partial\theta^2} l(\theta)\right] = \begin{bmatrix} \sum_i p_i(1-p_i) & \sum_i p_i(1-p_i)x_i \\ \sum_i p_i(1-p_i)x_i & \sum_i p_i(1-p_i)x_i^2 \end{bmatrix}$$

Now, $\hat{\theta}^0$ = MLE under $H_0: \beta_1 = 0$.

$$l^{H_0}(\beta_0,\beta_1) = l(\beta_0, 0) = \sum_i y_i \cdot \beta_0 - \log(1+\exp(\beta_0)) \Rightarrow \frac{\partial}{\partial \beta_0} l^{H_0} = \sum_i y_i - \frac{\exp(\beta_0)}{1+\exp(\beta_0)} \overset{set}{=} 0$$

$$\Rightarrow \frac{\exp(\hat\beta_0)}{1+\exp(\hat\beta_0)} = \frac{1}{n}\sum_i y_i$$

$$\Rightarrow \hat\beta_0 = \log\left(\frac{\bar y}{1-\bar y}\right), \quad \bar y := \frac{1}{n}\sum_i y_i$$

$$\therefore \hat\theta^0 = \begin{bmatrix} \text{logit}(\bar y) \\ 0 \end{bmatrix}$$

$$\Rightarrow p_i^0 = \frac{\exp(\hat\beta_0^0 + \hat\beta_1^0 x_i)}{1+\exp(\hat\beta_0^0 + \hat\beta_1^0 x_i)} = \frac{\exp(\text{logit}(\bar y))}{1+\exp(\text{logit}(\bar y))} = \bar y$$

$$\Rightarrow \dot{l}(\hat\theta^0) = \begin{bmatrix} \sum_i (y_i - \bar y) \\ \sum_i (y_i - \bar y)x_i \end{bmatrix} = \begin{bmatrix} 0 \\ \sum_i y_i x_i - n\bar y \bar x \end{bmatrix}, \quad I(\hat\theta^0) = \begin{bmatrix} \sum_i \bar y(1-\bar y) & \sum_i \bar y(1-\bar y)x_i \\ \sum_i \bar y(1-\bar y)x_i & \sum_i \bar y(1-\bar y)x_i^2 \end{bmatrix}$$

$$= n\bar y(1-\bar y)\begin{bmatrix} 1 & \bar x \\ \bar x & \overline{x^2} \end{bmatrix}$$

$$\Rightarrow I(\hat\theta^0)^{-1} = \frac{1}{n\bar y(1-\bar y)} \cdot \frac{1}{\overline{x^2} - \bar x^2}\begin{bmatrix} \overline{x^2} & -\bar x \\ -\bar x & 1 \end{bmatrix}$$

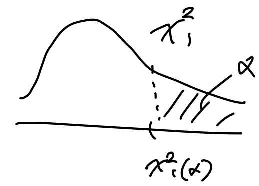
$$\therefore SC_n := \dot{l}(\theta)^T I(\theta)^{-1} \dot{l}(\theta)\Big|_{\theta = \hat\theta^0} \longrightarrow \chi^2_{2-1}$$

$$= \frac{1}{n\bar y(1-\bar y)\cdot(\overline{x^2} - \bar x^2)} \cdot \left(\sum_i y_i x_i - n\bar y \bar x\right)^2 \longrightarrow \chi^2_1$$

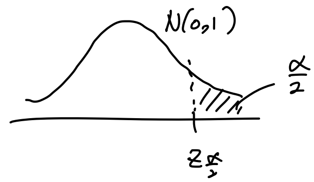
(c) cont'ed

Finally, Score test is given by

reject H_0 iff $$\frac{1}{n\bar{Y}(1-\bar{Y})\left(\overline{x^2}-\bar{x}^2\right)} \cdot \left(\sum_i x_i Y_i - \bar{x}\cdot \sum_i Y_i \right)^2 \ge \chi^2_1(\alpha)$$



or equivalenty $$\frac{\left| \sum_i x_i Y_i - \bar{x}\cdot \sum_i Y_i \right|}{\left(n\bar{Y}(1-\bar{Y}) \cdot \left(\overline{x^2}-\bar{x}^2\right)\right)^{1/2}} \ge z_{\frac{\alpha}{2}}$$



Comparing the approximate UMPU test derived in part (b):

reject H_0 iff $$\frac{\left| \sum_i x_i Y_i - \bar{x}\cdot \sum_i Y_i \right|}{\left(\frac{n}{n-1} \cdot n \cdot \bar{Y}(1-\bar{Y}) \cdot \left(\overline{x^2}-\bar{x}^2\right)\right)^{1/2}} \ge z_{\frac{\alpha}{2}}$$ //

→ Two are very similar except $\frac{n}{n-1}$ in the denominator of approximate UMPU test statistic.

(d)

Likelihood

$$L(\beta) = \prod_{i=1}^{n} \frac{\exp(y_i \cdot x_i^T\beta)}{1+\exp(x_i^T\beta)}$$

$$= \exp\left\{ \sum_i y_i \cdot x_i^T\beta - \sum_i \log(1+\exp(x_i^T\beta)) \right\}$$

$\underset{ff''(\beta)}{\underbrace{\overset{f(y)}{u} \cdot \overset{l^T\beta}{\theta} + \xi_1 \cdot T_1 + \cdots + \xi_{p-1} \cdot T_{p-1}}}$

$\sum y_i\, x_i^T \beta$

$(A^T)^{-1}A^T$

We can formulate an invertible matrix A having l as the first column.

$$\underset{p\times p}{A} = \left[l \,\vdots\, a_1 \,\vdots\, \cdots \,\vdots\, a_{p-1} \right] = \left[l \,\vdots\, \underset{p\times(p-1)}{A_0} \right]$$

$\underset{\text{matrix}}{A^T\beta} = \left[\overset{\theta}{l^T\beta} \,\vdots\, \underbrace{\cdots}_{p-1} \right]$, $\xi_1 \sim \xi_{p-1}$

Then, $x_i^T\beta = x_i^T(A^T)^{-1}A^T\beta = (A^{-1}x_i)^T(A^T\beta) = Z_i^T \begin{bmatrix} l^T\beta \\ a_1^T\beta \\ \vdots \\ a_{p-1}^T\beta \end{bmatrix} = Z_i^T \begin{bmatrix} \theta \\ \xi_1 \\ \vdots \\ \xi_{p-1} \end{bmatrix} = Z_{i0}\theta + Z_{i1}\xi_1 + \cdots + Z_{ip-1}\xi_{p-1}$

letting $Z_i := A^{-1}x_i = (z_{i0}, \cdots, z_{ip-1})^T$, $\theta = l^T\beta$, $\xi_j = a_j^T\beta$ $(1 \le j \le p-1)$

Then,

Likelihood

$$L(\theta, \xi_1, \cdots, \xi_{p-1}) = \exp\left\{ \sum_i y_i (Z_{i0}\theta + Z_{i1}\xi_1 + \cdots + Z_{ip-1}\xi_{p-1}) - \sum_i \log(1+\exp(z_{i0}\theta + z_{i1}\xi_1 + \cdots + z_{ip}\xi_{p-1})) \right\}$$

$$= \exp\left\{ \left(\sum_i z_{i0}y_i\right)\cdot\theta + \left(\sum_i z_{i1}y_i\right)\xi_1 + \cdots + \left(\sum_i z_{ip-1}\cdot y_i\right)\xi_{p-1} - \quad // \quad \right\}$$

Let $U := \sum_i z_{i0}Y_i$, $T_j = \sum_i z_{ij}Y_i$ $(1\le j \le p-1)$: ss for θ, ξ_j respectively.

Thus, UMPU size α test for testing $H_0: l^T\beta = \theta_0$ vs $H_1: l^T\beta \ne \theta_0$ is given by

(i) $\phi(Y) = \begin{cases} 1, & U < c_1(t) \text{ or } U > c_2(t) \\ \gamma_i, & U = c_i(t) \\ 0, & \text{o.w} \end{cases}$

(ii) $E_{H_0}[\phi(U) \mid T=t] = \alpha$

$E_{H_0}[U\phi(U) \mid T=t] = \alpha \cdot E_{H_0}[U \mid T=t]$.

for some $c_1(t), c_2(t)$: ftn of t and t is an arbitrary tuple of $T = (T_1, \cdots, T_{p-1})$.

To be specific, let obtain the cond'l distn of U given $T=t$ under H_0.

$Y_i \sim Ber(p_i)$, $p_i = \dfrac{\exp(x_i^T\beta)}{1+\exp(x_i^T\beta)} \overset{H_0}{=} \dfrac{\exp(z_{i0}\theta_0 + z_{i1}\xi_1 + \cdots + z_{ip-1}\xi_{p-1})}{1+\exp(\quad // \quad)}$

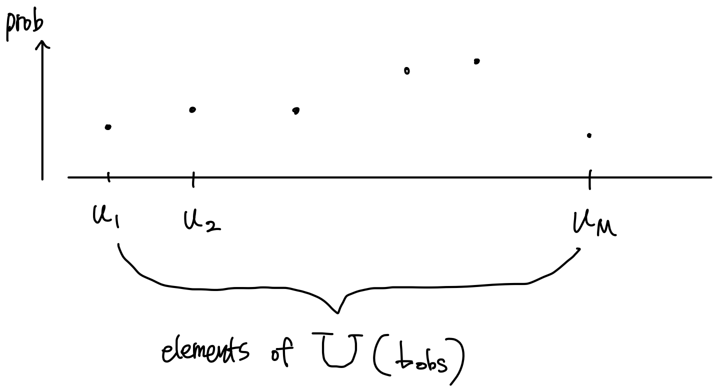
$$P_{H_0}(U=u \mid T=t) = \frac{P_{H_0}(U=u, T_1=t_1, \cdots, T_{p-1}=t_{p-1})}{P_{H_0}(\quad // \quad)} = \frac{|A(u, t_1, \cdots, t_{p-1})|}{\sum_{u' \in \underset{\sim}{U}(t)} |A(u', t_1, \cdots, t_{p-1})|}$$

where $A(u, t_1, \cdots, t_{p-1}) = \{ y : \sum_i z_{i0}y_i = u, \sum_i z_{i1}y_i = t_1, \cdots, \sum_i z_{ip-1}y_i = t_{p-1} \}$

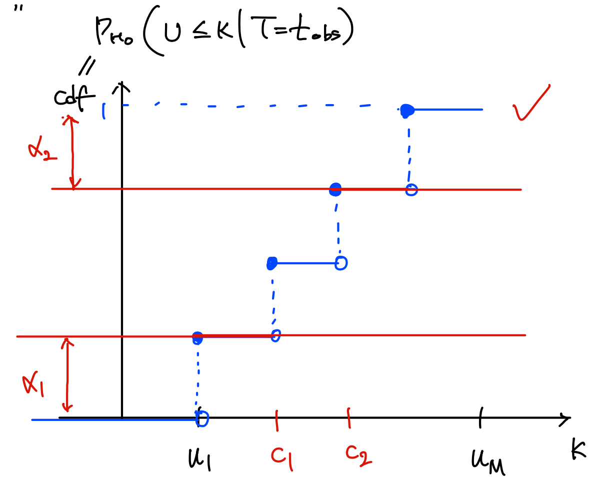
$\underset{\sim}{U}(t) = \{ u' \in \mathbb{R} : \sum_i z_{i0}y_i = u', \text{ where } y : \sum_i z_{i1}y_i = t_1, \cdots, \sum_i z_{ip-1}y_i = t_{p-1} \}$

set of all possible values of $\sum_i z_{i0}y_i$ where $T_j = t_j$.

(e) For the given data, we can compute u_{obs}, $t_{j,obs}$.
Then, we can get the dist^n of $U \mid T = t_{obs}$.



$\rightarrow$



For simplicity, let $f(u) = P_{H_0}(U = u \mid T = t_{obs})$
$F(u) = P_{H_0}(U \le u \mid T = t_{obs})$] we know both.

$U = \sum_{i=1}^{n} Z_i f(y_i)$, $0,1$, 2^n, "X" → "Z", "c_1" "c_2"

Then, we can select c_1 & c_2 : critical values s.t. $c_1, c_2 \in \mho(t_{obs})$,

$\alpha_1 = P_{H_0}(U < c_1 \mid T = t_{obs})$, $\alpha_2 = P_{H_0}(U > c_2 \mid T = t_{obs})$, $\alpha_1 + \alpha_2 < \alpha$, $\exists\, \gamma_1, \gamma_2 \in [0,1)$ s.t.

$$\alpha = \alpha_1 + \alpha_2 \cdot \gamma_1 \cdot P_{H_0}(U = c_1 \mid T = t_{obs}) + \gamma_2 \cdot P_{H_0}(U = c_2 \mid T = t_{obs})$$

or equivalently

$$\alpha = F(c_1) - f(c_1) + 1 - F(c_2) + \gamma_1 \cdot f(c_1) + \gamma_2 \cdot f(c_2)$$

and

$$\mathbb{E}_{H_0}[U\phi(U) \mid T = t_{obs}] = \alpha \cdot \mathbb{E}_{H_0}[U \mid T = t_{obs}]$$

or equivalently

$$\sum_{u < c_1} u \cdot f(u) + \sum_{u > c_2} u \cdot f(u) + \gamma_1 \cdot c_1 \cdot f(c_1) + \gamma_2 \cdot c_2 \cdot f(c_2) = \alpha \cdot \Big(\sum_{u \in \mho(t_{obs})} u \cdot f(u) \Big)$$

Since there are finite number of pairs of $(c_1, c_2) \in \mho(t_{obs})^2$, we can check every pair and conclude whether this pair satisfies above equations with some $\gamma_1, \gamma_2 \in [0,1)$.
We are guaranteed to find such pair by the UMPU Thm for exp'l family.

Now, we have fully formulated UMPU test for a given data. (i.e) $c_1, c_2, \gamma_1, \gamma_2$ found.

<NP Bootstrap Algorithm for computing the exact p-value> $X_1 \sim X_n \sim N(0,1)$ → $H_0: \mu = 0$ "B" "0.05 = α" small $|z| \ge c$ ~0.05

1. Data : (y_i, x_i) $1 \le i \le n$ $\Longrightarrow$ Convert into $X = (y_i, z_i)$ $1 \le i \le n$ by $z_i = A^{-1} x_i$.
2. For $b = 1, \cdots, B$, generate ^under H_0: bootstrapped sample $X_b^* = (y_{ib}^*, z_{ib}^*)$ by resampling (y_i, z_i) $1 \le i \le n$ with replacement
3. Perform UMPU test given in the above. That is, compute u_{obs}, t_{obs}, c_1, c_2 for X_b^*, and
   reject if $u_{obs} < c_1$ or $u_{obs} > c_2$ wp 1
   if $u_{obs} = c_1$ wp γ_1
   if $u_{obs} = c_2$ wp γ_2.
   "α"
4. Finally, exact p-value $= \dfrac{\#\text{ rejected tests}}{B}$

3.

(a) $S \sim B(n,p)$. $Y = (X_1, \dots, X_{S+1}) \mid S=s \sim N_s(\mu J_s, I_s)$

$$f_{S,Y}(s,y) = f_{Y|S}(y|s) \cdot f_S(s) = \left\{ \prod_{i=1}^{s+1} (2\pi)^{-1/2} \exp\left(-\tfrac{1}{2}(x_i-\mu)^2\right) \right\} \cdot \binom{n}{s} p^s (1-p)^{n-s}$$

$$= \exp\left\{ -\tfrac{1}{2}\sum x_i^2 + \mu \cdot \sum x_i - \tfrac{s+1}{2}\mu^2 + s \cdot \log\tfrac{p}{1-p} + n\log(1-p) \right\} \cdot (2\pi)^{-\frac{s+1}{2}} \cdot \binom{n}{s}$$

$$= \exp\left\{ \mu \cdot \sum x_i + \left(\log\tfrac{p}{1-p} - \tfrac{1}{2}\mu^2\right) \cdot s - \tfrac{1}{2}\mu^2 + n \cdot \log(1-p) \right\} \cdot c(s,y)$$

$\mathrm{supp}(f_{S,Y}) = \{0,1,\dots,n\} \times \mathbb{R}^{s+1}$: NOT depend on (p,μ) : parameters

$\theta := \log\frac{p}{1-p} - \frac{1}{2}\mu^2 \in \mathbb{R}$. $\Rightarrow (\mu,\theta) \in \mathbb{R}\times\mathbb{R}$: open set in $\mathbb{R}^2$

$f_{S,Y}$: full-rank exp'l family w/ rank $=2$

From the property of full-rank exp'l family, $(S, \sum_{i=1}^{S+1} X_i)$ is the CSS for (p,μ)

In case of $\mu = 0$,

$f_{S,Y}(s,y) = \exp\{ \log\frac{p}{1-p} \cdot s - B(p)\} \cdot c(s,y)$: full-rank exp'l family w/ rank $=1$

S : CSS for p

(b) From the property of full-rank exp'l family, (note that $(p,\mu) \mapsto (\theta,\mu)$ is 1-1)

$$(\hat{p},\hat{\mu}) = \arg_{(p,\mu)} \left[\mathbb{E}_{(p,\mu)}\left(\sum_{i=1}^{S+1} X_i, S\right) = \left(\sum_{i=1}^{S+1} X_i, S\right) \right]$$

Solving it,

$$\mathbb{E}_{(p,\mu)} \sum_{i=1}^{S+1} X_i = \mathbb{E}_{(p,\mu)}\left[\mathbb{E}_{(p,\mu)}\left(\sum_{i=1}^{S+1} X_i \middle| S\right)\right] = \mathbb{E}_{(p,\mu)}\left[(S+1)\mu\right] = (np+1)\mu$$

$$\mathbb{E}_{(p,\mu)} S = np$$

$$\therefore \begin{cases} (n\hat{p}+1)\hat{\mu} = \sum_{i=1}^{S+1} X_i \\ n\hat{p} = S \end{cases} \Rightarrow \begin{cases} \hat{p} = \frac{1}{n}S \\ \hat{\mu} = \frac{1}{S+1}\sum_{i=1}^{S+1} X_i \end{cases}$$

(c) $l(p,\mu) = \mu \sum x_i + \left(\log\frac{p}{1-p} - \frac{1}{2}\mu^2\right) \cdot s - \frac{1}{2}\mu^2 + n \cdot \log(1-p) + \log c(s,y)$

$$\frac{\partial l}{\partial p} = \left(\frac{1}{p} + \frac{1}{1-p}\right)s - \frac{n}{1-p}$$

$$\frac{\partial l}{\partial \mu} = \sum x_i - \mu s - \mu$$

$$\frac{\partial^2 l}{\partial p^2} = \left(-\frac{1}{p^2} + \frac{1}{(1-p)^2}\right)s - \frac{n}{(1-p)^2} \Rightarrow \mathbb{E}\left[\frac{\partial^2 l}{\partial p^2}\right] = \left(-\frac{1}{p^2} + \frac{1}{(1-p)^2}\right) \cdot np - \frac{n}{(1-p)^2} = -\frac{n}{p} + \frac{n(p-1)}{(1-p)^2} = -\frac{n}{p} - \frac{n}{1-p} = -\frac{n}{p(1-p)}$$

$$\frac{\partial^2 l}{\partial \mu \partial p} = 0 \qquad \Rightarrow \mathbb{E}\left[\frac{\partial^2 l}{\partial \mu \partial p}\right] = 0$$

$$\frac{\partial^2 l}{\partial \mu^2} = -S - 1 \qquad \Rightarrow \mathbb{E}\left[\frac{\partial^2 l}{\partial \mu^2}\right] = -np - 1$$

$$\therefore I(p,\mu) = \begin{bmatrix} \frac{n}{p(1-p)} & 0 \\ 0 & np+1 \end{bmatrix}$$

(c) cont'ed

$$I_1(p,\mu) = \lim_{n\to\infty} \frac{1}{n} I(p,\mu) = \begin{bmatrix} \frac{1}{p(1-p)} & 0 \\ 0 & p \end{bmatrix}$$

$$\therefore \sqrt{n}\left(\begin{bmatrix}\hat{p}\\ \hat{\mu}\end{bmatrix} - \begin{bmatrix}p\\ \mu\end{bmatrix}\right) \to_d N\left(\begin{bmatrix}0\\0\end{bmatrix}, \begin{bmatrix} p(1-p) & 0 \\ 0 & \frac{1}{p}\end{bmatrix}\right)$$

$$Var(\hat{p}) = \frac{1}{n^2} Var(S) = \frac{p(1-p)}{n}$$

(d) $H_0: \mu \le 0$ vs $H_1: \mu > 0$

ϕ : unbiased level α test of H_0 vs H_1

Unbiasedness : $\mathbb{E}_{\theta_0} \phi(S, X_1, \dots, X_{S+1}) \le \alpha$, $\forall \theta_0 \in \Theta_0 = \{(p,\mu) \in (0,1)\times(-\infty,0]\}$

$\mathbb{E}_{\theta_1} \phi(S, X_1, \dots, X_{S+1}) \ge \alpha$, $\forall \theta_1 \in \Theta_1 = \{(p,\mu) \in (0,1)\times(0,\infty)\}$

In terms of the power ft^n $\beta(p,\mu)$,

① $\sup_{p\in(0,1),\mu\le 0} \beta(p,\mu) \le \alpha$

② $\inf_{p\in(0,1),\mu>0} \beta(p,\mu) \ge \alpha$

To show $\beta(p,0) = \alpha$, $\forall p$,

we have $\beta(p,0) \le \alpha$, $\forall p$ from ①.

$$\mathbb{E}_{p,\mu}\, \phi(S, X_1, \dots, X_{S+1}) = \int \phi(s, x_1, \dots, x_{s+1}) \cdot \exp\left\{\mu \textstyle\sum x_i + \left(\log\frac{p}{1-p} - \frac{1}{2}\mu^2\right)\cdot s - \frac{1}{2}\mu^2 + n\log(1-p)\right\} \cdot c(s, x_1, \dots, x_{s+1})\, ds\, dx_1 \cdots dx_{s+1}$$

$\in [0,1]$: finitely exists

$\Rightarrow \beta(p,\mu)$: continuous in μ.

$\therefore \beta(p,\mu) \ge \alpha, \mu > 0$ & $\beta(p,\mu) \le \alpha, \mu \le 0$ implies that $\beta(p,0) = \alpha$.

(e) $f_{S,Y}(s,y) = f_{Y|S}(y|s) \cdot f_S(s)$

$$= \left(\prod_{i=1}^{s+1} (2\pi)^{-1/2} \exp\left(-\frac{1}{2}(x_i - \mu)^2\right)\right) \cdot \binom{n}{s} p^s (1-p)^{n-s}$$

$$= (2\pi)^{-(s+1)/2} \exp\left(-\frac{1}{2}\sum_{i=1}^{s+1} x_i^2 + \sum_{i=1}^{s+1} x_i \cdot \mu - \frac{(s+1)}{2}\mu^2 + s\log p + (n-s)\log(1-p)\right) \cdot \binom{n}{s}$$

$$= \exp\left(\mu \cdot \underbrace{\sum_{i=1}^{s+1} x_i}_{U} + \underbrace{\left(-\frac{\mu^2}{2} + \log\frac{p}{1-p}\right)}_{\eta} \cdot \underbrace{s}_{T} - \frac{1}{2}\mu^2 + n\log(1-p)\right) \cdot c(s,y)$$

$\eta = -\frac{\mu^2}{2} + \log\frac{p}{1-p}$: nuisance parameter.

UMPU of $H_0: \mu \le 0$ vs $H_1: \mu > 0$ is given by $\phi(S,Y) = \begin{cases} 1, & U > c(t) \\ \gamma, & U = c(t) \\ 0, & U < c(t) \end{cases}$, $\mathbb{E}_{\mu=0}[\phi(S,Y) | T = t] = \alpha$, where t : realization of T

$$P(U \le c \mid T = t) = P\left(\sum_{i=1}^{S+1} X_i \le c \mid S = t\right) = P\left(\sum_{i=1}^{t+1} X_i \le c\right) \qquad \boxed{\sum_{i=1}^{t+1} X_i \sim N(\mu(t+1), t+1)}$$

$$= P\left(\frac{\sum_{i=1}^{t+1} X_i - \mu(t+1)}{\sqrt{t+1}} \le \frac{c - \mu(t+1)}{\sqrt{t+1}}\right)$$

$$= \Phi\left(\frac{c - \mu(t+1)}{\sqrt{t+1}}\right)$$

$$\mathbb{E}_{\mu=0}\left[U \le c(t) \mid T = t\right] = \Phi\left(\frac{c(t)}{\sqrt{t+1}}\right) = \alpha.$$

$$\Rightarrow c(t) = Z_\alpha \cdot \sqrt{t+1}$$

$\therefore$ UMPU α test is given by

$$\phi(x) = I\left(U \le Z_\alpha \cdot \sqrt{T+1}\right) = I\left(\sum_{i=1}^{S+1} X_i \le Z_\alpha \cdot \sqrt{S+1}\right)$$