Statistics 582, Final Exam Solutions

Wellner; 3/16/2010

1. (30 points) **Define** *any three* of the following terms. In each case, provide an appropriate context for your definition.

   (a) An *unbiased test* of $H : \theta \in \Theta_0$ versus $K : \theta \in \Theta_1$.

   (b) A *similar on the boundary test* of $H : \theta \in \Theta_0$ versus $K : \theta \in \Theta_1$.

   (c) A *uniformly most powerful unbiased* level α test.

   (d) A family of distributions $\mathcal{P} = \{P_\theta : \ \theta \in \Theta \subset \mathbb{R}\}$ (with densities $p_\theta = dP_\theta/d\mu$ with respect to a dominating measure μ) with *monotone likelihood ratio.*

   (e) A *test with Neyman structure* with respect to a statistic T sufficient for $\mathcal{P}_B \equiv \{P_\theta : \ \theta \in \Theta_B\}$.

   Solution: See class notes, Chapter 6.

2. (30 points) **State** *any three* of the following results:

   (a) A theorem about admissibility properties of the sample mean $\overline{X}$ when sampling from a normal distribution on $\mathbb{R}$ and a contrasting theorem for sampling from a normal distribution on $\mathbb{R}^d$ with $d \geq 3$.

   (b) Stein's identity for $Eg'(X)$ where $X \sim N_1(0, \sigma^2)$ and $E|g'(X)| < \infty$.

   (c) A theorem relating Bayes rules to minimax rules and least favorable prior distributions.

   (d) The Wald-Wolfowitz-Noether-Hájek finite sampling central limit theorem.

   (e) The generalized Neyman - Pearson lemma (in the "short form" stated in the notes).

   Solution: See class notes, Chapters 5 and 6.

   Do one of problems 3, 4, and 5.

3. (40 points) Suppose that X has distribution function F on $[0, \infty)$.
   (a) Define the cumulative hazard function Λ in terms of F.
   (b) Give a general expression for $1-F$ in terms of the cumulative hazard function Λ, and show that it holds in the two cases (i) F is continuous; (ii) F is discrete with jumps $\Delta F_j \equiv F(x_j) - F(x_j-)$ at $x_1, \ldots, x_J$.

   Solution: See class notes, Chapter 4.

4. (40 points) **State** and **prove** the short form of the generalized Neyman - Pearson lemma.

   Solution: See class notes, Chapter 6.

5. (40 points) (a) Suppose that $Z \sim N(0,1)$ and $h : \mathbb{R} \to \mathbb{R}$ is a differentiable function with $E|h'(Z)| < \infty$. Give a heuristic proof of the identity $E(Zh(Z)) = Eh'(Z)$ using integration by parts.
   (b) Suppose that $X \sim N(\theta, \sigma^2)$ and $g : \mathbb{R} \to \mathbb{R}$ is a differentiable function satisfying $E|g'(X)| < \infty$. Use the result of (a) to prove Stein's identity.

   Solution: See class notes, Chapter 5.

   Do **either** problem 6 **or** problem 7.

6. (40 points) Consider the Beta(a,b) family of densities given by

$$p(x;a,b) = \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} x^{a-1}(1-x)^{b-1} 1_{(0,1)}(x), \qquad a,b > 0. \tag{1}$$

   Suppose that $X \sim p(\cdot; a, b)$. Consider testing:

$$\begin{aligned} H_1 &: \ a = b = 1 \ \text{ versus } \ K_1 : \ a = b = 2; \\ H_2 &: \ a = b \text{ (unspecified)versus } \ K_2 : \ b > a. \end{aligned}$$

   (a) Is the family of densities (1) an exponential family? Why or why not?
   (b) For $n = 1$ (or just one X with the distribution given in (1), what is the most powerful test of H_1 versus K_1 of size $\alpha = .1$?
   (c) For $n = 1$ (or just one X with the distribution given in (1), what theory do we have to yield tests of H_2 versus K_2 which are unbiased? Construct a uniformly most powerful test of H_2 versus K_2 of size $\alpha = .2$.
   (d) Now suppose that $X_1, \ldots, X_n$ i.i.d. with density given by (1). Give a reasonable test of H_2 versus K_2 based on $X_1, \ldots, X_n$.

   Solution:
   (a) Yes, since

$$p(x;a,b) = c(a,b)\exp(a\log x + b\log(1-x)) \cdot [x(1-x)]^{-1}$$

   with $c(a,b) \equiv \Gamma(a+b)/(\Gamma(a)\Gamma(b))$.
   (c) By the Neyman - Pearson lemma the most powerful test of size $\alpha = .1$ is given by

$$\phi(x) = \begin{cases} 1, & \text{if } \frac{p(x;2,2)}{p(x;1,1)} > c_\alpha \\ \gamma, & \text{if } \frac{p(x;2,2)}{p(x;1,1)} = c_\alpha \\ 0, & \text{if } \frac{p(x;2,2)}{p(x;1,1)} < c_\alpha \end{cases}$$

where c and γ are chosen so that $E_{(1,1)}\phi(X) = .1$. Now

$$\frac{p(x;2,2)}{p(x;1,1)} = \frac{\Gamma(4)}{\Gamma(2)^3} x(1-x) > c_\alpha$$

if and only if $x(1-x) > c'_\alpha$, and this occurs if and only if $x \in [1/2 - d, 1/2 + d]$ for some $d = d_\alpha$. Since

$$P_{(1,1)}(X \in [1/2 - d, 1/2 + d]) = 2d = \alpha = .1$$

if $d = .05$, the most powerful test of H_1 versus K_1 (based on one observation X) rejects H_1 if $.45 \le X \le .55$.
(c) Now

$$\begin{aligned} p(x;a,b) &= c(a,b)\exp((b-a)\log(1-x) + a\log(1-x) + a\log x) \\ &= c(a,b)\exp(\theta U(x) + \xi T(x)) \end{aligned}$$

with respect to the dominating measure ν determined by $\nu(A) = \int_A x^{-1}(1-x)^{-1}dx$; here $\theta \equiv b - a$, $U(x) \equiv \log(1-x)$, $\xi \equiv a$, and $T(x) \equiv \log[x(1-x)]$. Hence from the theory of unbiased tests, the UMP unbiased test of H_2 versus K_2 rejects when

$$\phi(X) = \begin{cases} 1, & \text{if } U(X) > c(T), \\ \gamma(T), & \text{if } U(X) = c(T), \\ 0, & \text{if } U(X) < c(T), \end{cases}$$

where $\gamma(T)$ and $c(T)$ are determined so that $E_{\theta=0}(\phi(X)|T) = \alpha$. To implement this test we need the conditional distribution of $U(X) = \log(1-X)$ conditional on $T(X) = \log[X(1-X)]$ under the null hypothesis. Now

$$\begin{aligned} &T(X) = \log[X(1-X)] = t \qquad \text{if and only if } \ X(1-X) = e^t \le 1/4, \\ &\text{if and only if } \ X^2 - X + e^t = 0, \qquad \text{if and only if } \ X = 1/2 \pm \sqrt{1/4 - e^t}, \end{aligned}$$

and under $a = b$ it is clear by symmetry that

$$P_{(a,a)}(X = 1/2 \pm \sqrt{1/4 - e^t}|T = t) = 1/2.$$

Since $U(X) > c(T)$ is equivalent to $X < 1 - e^{c(T)}$, the UMP unbiased test of H_2 versus K_2 of size $\alpha = .2$ becomes

$$\phi(X) = \begin{cases} .4, & \text{if } X = 1/2 - \sqrt{1/4 - e^T}, \\ 0, & \text{if } X = 1/2 + \sqrt{1/4 - e^T}. \end{cases}$$

Since $E_{(a,b)}X = a/(a+b) = 1/(1+b/a)$, this makes good sense intuitively.
(d) Two reasonable tests of $H_3 : a = b$ versus $K_3 :\ a \neq b$ could be based on the likelihood ratio and Rao statistics $2\log\lambda_n$ and R_n: "reject H_3 if

$$\left\{ \begin{array}{c} 2\log\lambda_n \\ R_n \end{array} \right\} > \chi^2_{1,\alpha}$$

where

$$\lambda_n \equiv \frac{\sup_{a>0,b>0} L_n(a,b)}{\sup_{a>0} L_n(a,a)} = \frac{L_n(\hat{a},\hat{b})}{L_n(\hat{a}_n^0,\hat{a}_n^0)}$$

where the likelihood $L_n(a,b)$ is given by

$$L_n(a,b) = c(a,b)^n \exp(a\sum_1^n \log X_i + b\sum_1^n \log(1-X_i)) \prod_1^n \{X_i(1-X_i)\}^{-1},$$

and where

$$R_n = Z_n^T I(\hat{a}_n^0,\hat{a}_n^0)^{-1} Z_n$$

with

$$Z_n \equiv \frac{1}{\sqrt{n}} \dot{\mathbf{l}}_n(\hat{\theta}_n^0).$$

7. (40 points) Suppose that X is a random variable with density $p(\cdot;\theta)$ given by

$$p(x;\theta) = \frac{1}{\pi}\frac{1}{\sqrt{x(1-x)}} \cdot \frac{\theta}{1-(1-\theta^2)x} 1_{(0,1)}(x), \quad \theta \in (0,\infty). \tag{2}$$

This is the distribution of the total time spent in $(0,\infty)$ up to time $t = 1$ by a skew Brownian motion process with skewing parameter $\theta = \sigma_+/\sigma_-$ where σ_+^2 is the variance parameter for the positive space axis and σ_-^2 is the variance parameter for the negative space axis. Note that for $\theta = 1$,

$$p(x,1) = \frac{1}{\pi}\frac{1}{\sqrt{x(1-x)}} 1_{(0,1)}(x)$$

is the Beta$(1/2,1/2)$ density corresponding to the arcsin distribution

$$F_1(x) = P_1(X \le x) = \frac{2}{\pi}\arcsin(\sqrt{x}).$$

The mean and variance of X with density $p(\cdot;\theta)$ are $E_\theta X = 1/(1+\theta)$ and $Var_\theta(X) = 2^{-1}\frac{1}{1+\theta}\cdot\frac{\theta}{1+\theta}$.
(a) Show that the family $\mathcal{P} = \{p(\cdot;\theta) :\ \theta \in (0,\infty)\}$ has monotone likelihood ratio in $T(x) = 1/x$. [Hint: rewrite the last factor in (2) in terms of $1/x$.]
(b) Find the UMP size .05 test of $H :\ \theta \le 1 \equiv \theta_0$ versus $K :\ \theta > 1$? Specify

your test completely, including the constant(s).
(c) Compute the power function of the UMP test in (b) as explicitly as possible.
(d) Now suppose that $X_1, \ldots, X_n$ are i.i.d. with density $p(\cdot;\theta)$. Find the form of the locally most powerful test of $H: \theta \le 1$ versus $K: \theta > 1$ based on $X_1, \ldots, X_n$, and use the central limit theorem to find appropriate constants so that your test has approximate size $\alpha = .05$.

Solution: (a) Suppose that $0 < \theta < \theta' < \infty$. Then

$$\begin{aligned} \frac{p(x;\theta')}{p(x;\theta)} &= \frac{\theta'}{1-(1-\theta'^2)x} \cdot \frac{1-(1-\theta^2)x}{\theta} = \frac{\theta'}{\theta} \cdot \frac{(1/x)-(1-\theta^2)}{(1/x)-(1-\theta'^2)} \\ &\equiv \frac{\theta'}{\theta} \cdot \frac{t-(1-\theta^2)}{t-(1-\theta'^2)}. \end{aligned}$$

Thus

$$\begin{aligned} \frac{d}{dt} \log\left(\frac{p(1/t;\theta')}{p(1/t;\theta)}\right) &= \frac{1}{t-(1-\theta^2)} - \frac{1}{t-(1-\theta'^2)} \\ &= \frac{\theta'^2-\theta^2}{(t-(1-\theta^2))(t-(1-\theta^2))} \\ &> 0. \end{aligned}$$

Thus the family $\mathcal{P} = \{p(\cdot;\theta) : \theta > 0\}$ has MLE in $T(x) = 1/x$.
(b) It follows from the Karlin-Rubin theorem that the UMP size α test of $H: \theta \le 1$ versus $K: \theta > 1$ is of the form $\phi(X) = 1\{1/X > c\} + \gamma 1\{1/X = c\} = 1\{X < 1/c\} + \gamma 1\{X = 1/c\}$ where c and γ are determined by $\alpha = E_{\theta_0=1}\phi(X) = P_1(X < 1/c) + \gamma P_1(X = 1/c)$. Since the distribution of X is continuous we can take $\gamma = 0$ and choose $1/c \equiv \tilde{c}$ to satisfy $P_1(X < \tilde{c}) = \alpha = .05$. But since we know that $P_1(X \le x) = (2/\pi)\arcsin(\sqrt{x})$, this yields $\tilde{c} = \tilde{c}_\alpha = (\sin(\pi\alpha/2))^2 = (\sin(\pi/40))^2$.
(c) The power function of the test in (b) is just

$$\begin{aligned} \beta_\phi(\theta) &= E_\theta\phi(X) = P_\theta(X < \tilde{c}_\alpha) = \int_0^{\tilde{c}_\alpha} p(x;\theta)dx \\ &= \int_0^{\tilde{c}_\alpha} \frac{1}{\pi}\frac{1}{\sqrt{x(1-x)}} \cdot \frac{\theta}{1-(1-\theta^2)x} dx \\ &= 1 - \frac{2}{\pi}\arctan\left(\frac{\cot(\alpha\pi/2)}{\theta}\right) \end{aligned}$$

after three changes of variables: $x = t^2$; followed by $t = 1/y$; followed by $v = \sqrt{y^2-1}$. This last formula is due to Anna Klimova, who computed it during the exam. See Figure 1 for a plot of this power function.

Here is a detailed derivation of the power formula: by letting $x = t^2$ we find that

$$
\begin{aligned}
&\int_0^{\tilde{c}_\alpha} \frac{1}{\pi} \frac{1}{\sqrt{x(1-x)}} \cdot \frac{\theta}{1-(1-\theta^2)x} dx \\
= & \int_0^{\sin(\alpha\pi/2} \frac{1}{\pi} \frac{\theta 2t}{\sqrt{t^2-t^4} \cdot (1-(1-\theta^2)t^2)} dt \\
= & \int_0^{\sin(\alpha\pi/2} \frac{2}{\pi} \frac{\theta}{\sqrt{1-t^2} \cdot (1-(1-\theta^2)t^2} dt \\
= & \frac{2\theta}{\pi} \int_0^{\sin(\alpha\pi/2)} \frac{1}{t^2\sqrt{1-t^2} \cdot (1/t^2-(1-\theta^2))} dt \\
= & \frac{2\theta}{\pi} \int_0^{1/\sin(\alpha\pi/2)} \frac{-1/y^2}{(1/y^2)\sqrt{1-1/y^2} \cdot (y^2-(1-\theta^2)} dy, \quad t = 1/y, \\
= & \frac{2\theta}{\pi} \int_{1/\sin(\alpha\pi/2)}^{\infty} \frac{y}{\sqrt{y^2-1} \cdot (y^2-1+\theta^2)} dy \\
= & \frac{2\theta}{\pi} \int_{\cot(\alpha\pi/2)}^{\infty} \frac{1}{v^2+\theta^2} dv, \quad v = \sqrt{y^2-1} \\
= & \frac{2\theta}{\pi} \frac{1}{\theta} \arctan(v/\theta) \Big|_{\cot(\alpha\pi/2)}^{\infty} \\
= & \frac{2}{\pi} \left(\frac{\pi}{2} - \arctan\left(\frac{\cot(\alpha\pi/2)}{\theta} \right) \right) \\
= & 1 - \frac{2}{\pi} \arctan\left(\frac{\cot(\alpha\pi/2)}{\theta} \right).
\end{aligned}
$$

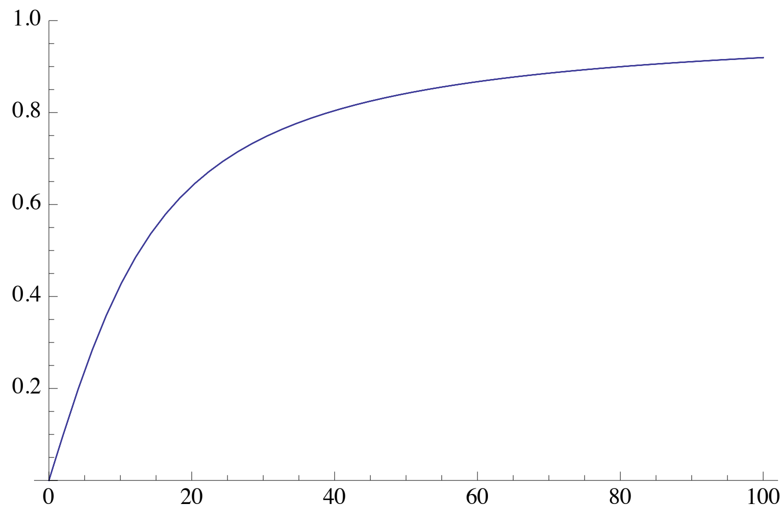

Figure 1: Plot of power, UMP test for $n = 1$.

(d) As we have see in the context of the Cauchy location family, the locally most powerful test of $H : \theta = \theta_0 = 1$ versus $K : \theta > 1$ is of the form "reject H if $\dot{\mathbf{l}}_{n\theta}(\underline{X}; \theta_0) > k$" where $\dot{\mathbf{l}}_{n,\theta}$ is the score function based on all the data and k is a constant to be determined so that the test has (at least approximately) size α. In this present case, the score for one observation is

$$\begin{aligned} \dot{\mathbf{l}}_\theta(x; \theta) &= \frac{1}{\theta} - \frac{2\theta x}{1-(1-\theta^2)x} = \frac{1-(1+2\theta-\theta^2)x}{\theta(1-(1-\theta^2)x} \\ &= 1 - 2x \qquad \text{when} \quad \theta = \theta_0 = 1. \end{aligned}$$

Thus the locally most powerful test of $H : \theta = 1$ versus $K : \theta > 1$ is of the form "reject $H : \theta = 1$ if $\sum_1^n (1 - 2X_i) > k$; or, equivalently when $\overline{X}_n - (1/2) < k'$; or, equivalently, when $Z_n \equiv \sqrt{n}(\overline{X}_n - 1/2)/\sqrt{1/8} < k''$. Under $\theta_0 = 1$ the X_i's have $E_1(X_1) = 1/2$ and $Var_1(X_1) = 1/8$, and therefore the CLT implies that the test

$$\phi(\underline{X}) = 1\{Z_n < -z_\alpha\}, \quad \text{where } z_\alpha \equiv \Phi^{-1}(1-\alpha)$$

has $E_1 \phi(\underline{X}) \to \alpha$.

Do **either** problem 8 **or** problem 9.

8. (40 points). Suppose that conditional on $\boldsymbol{\theta} = \theta$, $X \sim$ Uniform$(0, \theta)$ and that $\boldsymbol{\theta} \sim$ Gamma$(2, \mu)$; $\boldsymbol{\theta}$ has density $\lambda(\theta) = \mu^2 \theta exp(-\mu\theta) 1_{(0,\infty)}(\theta)$ for $\theta > 0$.

(a) Find the marginal density of X and the conditional density of $\boldsymbol{\theta}$ given $X = x$.
(b) What is the Bayes estimator of θ (for squared error loss), $d_\mu(X) \equiv d_{\Lambda_\mu}(X)$?
(c) Find the maximum likelihood estimator of μ based on the marginal distribution of X, and use this to find the empirical Bayes estimator $d_{EB}(X) \equiv d_{\hat{\mu}}(X)$ of θ.
(d) Compute and compare the risks of the maximum likelihood estimator d_{ML}, the Bayes estimator d_B, and the empirical Bayes estimator d_{EB}.

Solution: (a) The marginal density $q \equiv q_\mu$ of X is given by

$$\begin{aligned} q(x) &= \int_0^\infty \frac{1}{\theta} 1_{(0,\theta)}(x) \mu^2 \theta \exp(-\mu\theta) d\theta \\ &= \mu \int_x^\infty \mu \exp(-\mu\theta) d\theta = \mu \exp(-\mu x); \end{aligned}$$

i.e. the marginal density of X is Exponential(μ). The conditional density of $\boldsymbol{\theta}$ given $X = x$ is given by

$$\begin{aligned} \lambda(\theta|x) &= \frac{\theta^{-1} 1_{(0,\theta)}(x) \mu^2 \theta \exp(-\mu\theta)}{\mu \exp(-\mu x)} \\ &= \mu \exp(-(\theta - x)) 1_{(x,\infty)}(\theta), \end{aligned}$$

i.e. the Exponential (μ) distribution shifted to the right by x.
(b) It follows immediately from (a) that the Bayes estimator with respect to squared error loss is given by $E(\boldsymbol{\theta}|X) = X + 1/\mu \equiv d_{\Lambda_\mu}(X) \equiv d_\mu(X)$.
(c) Now from (a) we find that the maximum likelihood estimator of μ based on the marginal density of X is $\hat{\mu} = 1/X$. Plugging this into the Bayes estimator we found in (b) yields the empirical Bayes estimator

$$d_{EB}(X) = d_{\hat{\mu}}(X) = X + 1/\hat{\mu} = 2X.$$

(d) The MLE of θ based on X is $d_{ML}(X) = X$. This estimator has risk

$$\begin{aligned} R(\theta, d_{ML}) &= Var_\theta(X) + (E_\theta(X) - \theta)^2 \\ &= \frac{\theta^2}{12} + \left(\frac{\theta}{2}\right)^2 = \frac{\theta^2}{3}. \end{aligned}$$

The Bayes estimator has risk $R(\theta, d_B)$ given by

$$\begin{aligned} R(\theta, d_B) &= Var_\theta(X + \mu^{-1}) + (E_\theta(X + \mu^{-1} - \theta)^2 \\ &= \frac{\theta^2}{12} + (\mu^{-1} - \theta/2)^2 \geq \frac{\theta^2}{12} \end{aligned}$$

with equality if and only if $\mu^{-1} = \theta/2$. Finally the empirical Bayes estimator d_{EB} has risk $R(\theta, d_{EB})$ given by

$$R(\theta, d_{EB}) = Var_\theta(2X) + (E_\theta(2X) - \theta)^2 = 4\frac{\theta^2}{12} = \frac{\theta^2}{3}.$$

Thus the risk of the empirical Bayes estimator is the same as that of the ML estimator: although the EB estimator has zero bias, it has larger variance. Although the risk of the Bayes estimator is smaller than either the ML or EB estimator if μ^{-1} is close to $\theta/2$, it has worse risk if $\mu^{-1} > \theta$.
Note: This problem generalizes nicely to $X_1, \ldots, X_n$ i.i.d. Uniform$(0, \theta)$ and $\boldsymbol{\theta} \sim$ Gamma$(n+1, \mu)$. Then the empirical Bayes estimator turns out to be the unbiased version of the MLE, namely $\hat{\theta}_{EB} = ((n+1)/n)X_{(n)}$ and $\hat{\theta}_{EB}$ has smaller risk than the MLE for $n > 1$ with a limiting risk ratio of $1/2$.

9. (40 points). Suppose that $X_1, \ldots, X_m$ are i.i.d. Poisson(μ) and $Y_1, \ldots, Y_n$ are i.i.d. Poisson (ν) random variables independent of the X_i's. Consider testing $H : \mu \ge \nu$ versus $K : \nu > \mu$.
(a) Find the joint distribution of $(\underline{X}, \underline{Y})$ and put it in exponential family form. What are the marginal distributions of $R \equiv \sum_1^m X_i$, $S \equiv \sum_1^n Y_j$, and $T = R+S = \sum_i X_i + \sum_j Y_j$? (b) Find the conditional distribution of $S = \sum_1^n Y_j$ given $T = t$ and specialize this to the case when $\mu = \nu$.
(c) Find the UMP unbiased test of H versus K of size α and describe how you would find the constants involved as explicitly as possible.
(d) Show that you can rewrite the test you derived in (c) as $\phi(\underline{X}, \underline{Y}) = 1\{V(\underline{X}, \underline{Y}) > c(T)\} + \gamma(T)1\{V(\underline{X}, \underline{Y}) = c(T)\}$ where

$$V(\underline{X}, \underline{Y}) \equiv \frac{\sum_{j=1}^n Y_j - T\frac{n}{N}}{\sqrt{T\frac{n}{N}\frac{m}{N}}} = \frac{\sqrt{\frac{mn}{N}}(\overline{Y}_n - \overline{X}_m)}{\sqrt{T/N}}. \tag{3}$$

(e) Use (d), the CLT, and the WLLN to show that if the constant in (d) is taken to be $z_\alpha \equiv \Phi^{-1}(1-\alpha)$, then the test in (d) has approximate size α.

Solution: (a) The joint mass function (or density with respect to counting measure on $\mathbb{N}^{m+n}$) of $(\underline{X}, \underline{Y})$ is given by

$$\begin{aligned}
p_{\mu,\nu}(\underline{x}, \underline{y}) &= \prod_{i=1}^{m} e^{-\mu}\frac{\mu^{x_i}}{x_i!} \cdot \prod_{j=1}^{n} e^{-\nu}\frac{\nu^{y_j}}{y_j!} \\
&= \exp(-m\mu - n\nu)\mu^{\sum x_i}\nu^{\sum y_j}\frac{1}{\prod x_i! \prod y_j!} \\
&= c_{m,n}(\mu,\nu)\exp\left(\sum_1^m x_i \log\mu + \sum_1^n y_j \log\nu\right)\frac{1}{\prod x_i! \prod y_j!} \\
&= c_{m,n}(\mu,\nu)\exp\left(\sum_1^n y_j \log(\nu/\mu) + (\sum_1^m x_i + \sum_1^n y_j)\log\mu\right)\frac{1}{\prod x_i! \prod y_j!} \\
&= \tilde{c}_{m,n}(\theta,\xi)\exp\left(\theta U(\underline{x},\underline{y}) + \xi T(\underline{x},\underline{y})\right)\frac{1}{\prod x_i! \prod y_j!}
\end{aligned}$$

where $U \equiv \sum_1^n Y_j$, $\theta \equiv \log(\nu/\mu)$, $T \equiv \sum_1^m X_i + \sum_1^n Y_j \equiv R + S$, $\xi \equiv \log\mu$. The marginal distributions of $R = \sum_1^m X_i$ and $S = \sum_1^n Y_j$ are just Poisson$(m\mu)$ and Poisson $(n\nu)$ respectively, while the sum $T = R + S$ is marginally Poisson $(m\mu + n\nu)$.
(b) The conditional distribution of U given T is

$$\begin{aligned}
P_{\mu,\nu}(S = s|T = t) &= \frac{P_{\mu,\nu}(S = s, R = t - s)}{P_{\mu,\nu}(T = t)} \\
&= \frac{e^{-n\nu}\frac{(n\nu)^s}{s!} \cdot e^{-m\mu}\frac{(m\mu)^{t-s}}{(t-s)!}}{\exp(-(m\mu + n\nu))\frac{(m\mu+n\nu)^t}{t!}} \\
&= \binom{t}{s}\left(\frac{n\nu}{m\mu + n\nu}\right)^s\left(\frac{m\mu}{m\mu + n\nu}\right)^{t-s}.
\end{aligned}$$

Thus on the boundary $\Theta_B = \{(\mu,\mu): \ \mu > 0\}$, $(S|T = t) \sim$ Binomial$(t, \frac{n}{m+n})$.
(c) The UMP unbiased set of size α is of the form

$$\phi(\underline{X}, \underline{Y}) = 1\{U > c(T)\} + \gamma(T)1\{U = c(T)\}$$

where $c(t)$ and $\gamma(t)$ are determined by

$$\begin{aligned}
\alpha &= E_{\mu,\mu}\{\phi(\underline{X}, \underline{Y})|T\} \\
&= P(\text{Bin}(T, n/(m+n)) > c(T)) + \gamma(T)P(\text{Bin}(T, n/(m+n)) = c(T)).
\end{aligned}$$

(d) Now the test we derived in (b) and (c) is, up to randomization on the boundary

$c(t)$,

$$
\begin{aligned}
&\text{“reject if } \sum_{j=1}^{m} Y_j > c(T)\text{”}, \qquad \text{or equivalently} \\
&\text{“reject if } \sum_{j=1}^{m} Y_j - \frac{n}{N} T > c_1(T) \equiv c(T) - (n/N)T\text{”}, \qquad \text{or equivalently} \\
&\text{“reject if } \sum_{j=1}^{m} Y_j - \frac{n}{N}\sum_{j=1}^{m} Y_j - \frac{n}{N}\sum_{i=1}^{n} X_i > c_1(T)\text{”}, \qquad \text{or equivalently} \\
&\text{“reject if } \frac{m}{N}\sum_{j=1}^{m} Y_j - \frac{n}{N}\sum_{i=1}^{n} X_i > c_1(T)\text{”}, \qquad \text{or equivalently} \\
&\text{“reject if } \frac{mn}{N}(\overline{Y} - \overline{X}) > c_1(T)\text{”}, \qquad \text{or equivalently} \\
&\text{“reject if } \frac{\frac{mn}{N}(\overline{Y} - \overline{X})}{\sqrt{T\frac{n}{N}\frac{m}{N}}} > c_2(T) \equiv \frac{c_1(T)}{\sqrt{T\frac{n}{N}\frac{m}{N}}}\text{”}, \qquad \text{or equivalently} \\
&\text{“reject if } \frac{\sqrt{\frac{mn}{N}}(\overline{Y} - \overline{X})}{\sqrt{\frac{T}{N}}} > c_2(T)\text{”}.
\end{aligned}
$$

in other words, the conditional test rejects H when $V(\underline{X}, \underline{Y}) > \tilde{c}(T)$ where $V(\underline{X}, \underline{Y})$ is as given in (3).
(e) Now when $\mu = \nu$ is true, if we assume that $m/N \equiv \lambda_N \to \lambda$, then by the CLT

$$
\begin{aligned}
\sqrt{mn}N(\overline{Y} - \overline{X}) &= \sqrt{n/N}\sqrt{m}(\overline{Y} - \nu) - \sqrt{m/N}\sqrt{n}(\overline{X} - \nu) \\
&\to_d \sqrt{1-\lambda}N(0,\nu) - \sqrt{\lambda}N(0,\nu) \stackrel{d}{=} N(0,\nu),
\end{aligned}
$$

and $T/N = \lambda_N \overline{X} + (1-\lambda_N)\overline{Y} \to_p \lambda\nu + (1-\lambda)\nu = \nu$. Thus it follows by Slutsky's theorem that $V(\underline{X}, \underline{Y}) \to_d Z \sim N(0,1)$. Hence if we choose $\tilde{c}(T) = z_\alpha \equiv \Phi^{-1}(1-\alpha)$, the resulting test “reject H if $V(\underline{X}, \underline{Y}) > z_\alpha$” has approximate size α for large m, n. (In fact this remains true even if λ_N does not converge, since the limits remain the same along any convergent subsequence, and convergent subsequences can be selected since λ_N takes values in the compact set $[0, 1]$.)