Statistics 582, Problem Set 7 Solutions

Wellner; 2/25/2009

1. Continuation of problem 3, problem set 4: Suppose that $X_1, \ldots, X_n$ are i.i.d. Exponential(θ) (so the X's have distribution P_θ and density $p_\theta(x) = \theta e^{-\theta x} 1_{(0,\infty)}(x)$. with respect to Lebesgue measure on $\mathbb{R}$, and that $\theta \sim \Gamma(\alpha, \beta)$:

$$\lambda(\theta) = \beta \frac{(\beta\theta)^{\alpha-1}}{\Gamma(\alpha)} \exp(-\beta\theta) 1_{[0,\infty)}(\theta).$$

In problem set 4 we found the Bayes rules with respect to squared error loss $L(\theta, a) = (\theta - a)^2$ and weighted squared error loss $L(\theta, a) = (\theta - a)^2/\theta$.

(a) Prove a (conditional) limit theorem for the posterior distributions given $\underline{X}$.
(b) What does theorem 5.8.2 say about the limiting distribution of the Bayes rule for squared error loss (assuming that $X_1, \ldots, X_n$ are i.i.d. $P_{\theta_0} \equiv P$ with $\theta_0 \in (0, \infty)$)?

Solution: (a) There are several possible ways of proceeding here: (i) verify the hypotheses of Theorem 8.1 of the notes; (ii) verify the hypotheses of Theorem 10.1 of van der Vaart's *Asymptotic Statistics*; (iii) give a direct proof in this special case of convergence in distribution; or (iv) give a direct proof in this special case of convergence in total variation distance by showing that the densities converge pointwise followed by Scheffé's lemma. Both (i) and (ii) are made difficult by conditions B2 and B3 (in the case of Theorem 8.1) and by the "separation by tests" condition in van der Vaart's Theorem 10.1. Thus proceed directly as in (iii). First, note that

$$\begin{aligned}\theta \sim \text{Gamma}(\alpha + n, \beta + \sum X_i) &=_d (\beta + \sum X_i)^{-1}\text{Gamma}(\alpha + n, 1) \\ &=_d (\beta + \sum X_i)^{-1}(Y_0 + \sum_{i=1}^n Y_i)\end{aligned}$$

where $Y_0 \sim \text{Gamma}(\alpha, 1)$, and $Y_i \sim \text{Gamma}(1,1) = \text{Exp}(1)$, $i = 1, \ldots, n$ are all independent. Thus conditionally on the X_i's we have, with $Z \sim N(0,1)$ and with θ_0 the true value of θ,

$$\begin{aligned}\sqrt{n}(\theta - E(\theta|\underline{X})) &=_d \sqrt{n}\frac{Y_0 + \sum_{i=1}^n Y_i - (\alpha + n)}{\beta + \sum_{i=1}^n X_i} \\ &= \sqrt{n}(\overline{Y}_n - 1)\frac{1}{\overline{X}_n + n^{-1}\beta} + \sqrt{n}(Y_0 - \alpha)\frac{1/n}{\overline{X}_n + n^{-1}\beta} \\ &\to_d Z\frac{1}{\theta_0^{-1}} \sim N(0, \theta_0^2)\end{aligned}$$

almost surely with respect to the distribution of $X_1, X_2, \ldots$. Note that the posterior mean $E(\theta|\underline{X})$ can be replaced here by either the MLE $1/\overline{X}_n$ or by $T_n = \theta_0 + (nI(\theta_0))^{-1}\sum_{i=1}^n \dot{l}_\theta(X_i) = 2\theta_0 - \theta_0^2\overline{X}_n$ since

$$\sqrt{n}(E(\theta|\underline{X}) - 1/\overline{X}_n) = o_p(1)$$

and similarly with T_n in place of $1/\overline{X}_n$.
To show that the densities of $\sqrt{n}(\theta - E(\theta|\underline{X})$ converge pointwise, first consider the distribution function and density of the unscaled version of the lead term, $\sqrt{n}(\overline{Y}_n - 1)$: since $n\overline{Y}_n = \sum_1^n Y_i \sim \text{Gamma}(n, 1)$,

$$\begin{aligned} F_{\sqrt{n}(\overline{Y}_n - 1)}(z) &= P(\overline{Y}_n \le 1 + n^{-1/2}z) = P(\sum_1^n Y_i \le n + \sqrt{n}z) \\ &= \int_0^{n+\sqrt{n}z} \frac{t^{n-1}}{\Gamma(n)} \exp(-t)dt. \end{aligned}$$

Thus, differentiating and then using

$$\Gamma(n) = (n-1)! \sim \sqrt{2\pi(n-1)}\left(\frac{n-1}{e}\right)^{n-1}$$

by Stirling's formula, we find that

$$\begin{aligned} \sqrt{n}(\overline{Y}_n - 1) &= \frac{(n+\sqrt{n}z)^{n-1}}{\Gamma(n)} \exp(-(n+\sqrt{n}z))\sqrt{n} \\ &\sim \frac{(n+\sqrt{n}z)^{n-1}}{\sqrt{2\pi(n-1)}\left(\frac{n-1}{e}\right)^{n-1}} \cdot \exp(-(n+\sqrt{n}z)) \cdot \sqrt{n} \\ &= \frac{1}{\sqrt{2\pi}}\sqrt{\frac{n}{n-1}} \cdot \left(1 + \frac{1+\sqrt{n}z}{n-1}\right)^{n-1} \cdot \exp(-(1+\sqrt{n}z)) \\ &\to \frac{1}{\sqrt{2\pi}} \exp(-z^2/2). \end{aligned}$$

Here the convergence follows by letting $a_n \equiv 1 + \sqrt{n+1}z$ and noting that

$$\begin{aligned}
&\left(1 + \frac{1+\sqrt{n+1}z}{n}\right)^n \cdot \exp(-(1+\sqrt{n+1}z)) \\
&= \left(1 + \frac{a_n}{n}\right)^n \cdot \exp(-a_n) \\
&= \left\{\left(1 + \frac{a_n}{n}\right) \cdot \exp(-a_n/n)\right\}^n \\
&\approx \left\{\left(1 + \frac{a_n}{n}\right)\left(1 - \frac{a_n}{n} + \frac{1}{2}\frac{a_n^2}{n} + O(n^{-3/2})\right)\right\}^n \\
&= \left\{1 - \frac{1}{2}\frac{a_n^2}{n^2} + O(n^{-3/2})\right\}^n \\
&= \left\{1 - \frac{1}{2}\frac{(1+\sqrt{n+1}z)^2/n}{n}\right\}^n \\
&\to \exp(-z^2/2).
\end{aligned}$$

Since $\overline{X}_n + n^{-1}\beta \to_{a.s.} \theta_0^{-1}$ and $\sqrt{n}(Y_0 - \alpha)\frac{1/n}{\overline{X}_n + n^{-1}\beta} \to_{a.s.} 0$, it follows (via the convolution formula) that the density of $\sqrt{n}(\theta - E(\theta|\underline{X})$ converges pointwise to $\phi(z/\theta_0)/\theta_0$, the density of $N(0, \theta_0^2)$.

(b) In the present case Theorem 5.8.2 says that

$$\sqrt{n}(E(\theta|\underline{X}) - \theta_0) \to_d N(0, 1/I(\theta_0)) = N(0, \theta_0^2)$$

since $I(\theta_0) = 1/\theta_0^2$. This also follows from a direct argument since

$$\begin{aligned}
\sqrt{n}(E(\theta|\underline{X}) - \theta_0) &= \sqrt{n}\left(\frac{1+\alpha/n}{\overline{X}_n + \beta/n} - \theta_0\right) \\
&= \sqrt{n}(\alpha/n + 1 - \theta_0(\beta/n + \overline{X}_n))/(\beta/n + \overline{X}_n) \\
&= \left\{-\theta_0\sqrt{n}(\overline{X}_n - 1/\theta_0) + n^{-1/2}(\alpha - \theta_0\beta)\right\}/(\beta_n + \overline{X}_n) \\
&\to_d \theta_0^2 N(0, \theta_0^{-2}) = N(0, \theta_0^2).
\end{aligned}$$

2. A random variable X takes on the values 1, 2, 3, 4 with probability distribution $p_0(x)$ or $p_1(x)$ as follows:

x	1	2	3	4
$p_0(x)$	.18	.06	.36	.40
$p_1(x)$	.36	.18	.24	.22

A. Find a most powerful test of size $\alpha = .2$ for testing p_0 versus p_1 and determine its power.
B. Find a test ϕ which minimizes the sum of risks $a + b$ where $a = E_0\phi$ and $b = E_1(1-\phi)$.

Solution: A. Now $p_1(x)/p_0(x) = 2,\ 3,\ 2/3,\ 22/40 = 11/20$, according as $x = 1,\ 2,\ 3,\ 4$, so a MP test of size $\alpha = .2$ is given by

$$\phi(x) = \begin{cases} 1, & \text{if } x = 2 \\ .14/.18, & \text{if } x = 1 \\ 0, & \text{if } x = 3, 4. \end{cases}$$

Then

$$E_0\phi(X) = P_0(X=2) + \frac{.14}{.18}P_0(X=1) = .06 + \frac{.14}{.18}.18 = .2,$$

while

$$\text{Power} = E_1\phi(X) = P_1(X=2) + \frac{.14}{.18}P_1(X=1) = .18 + \frac{.14}{.18}.36 = .18 + .28 = .46.$$

B. A test which minimizes $a + b$ is given by

$$\begin{aligned} \phi(x) &= \begin{cases} 1, & \text{if } \ p_1(x) \ge p_0(x) \\ 0, & \text{if } \ p_1(x) < p_0(x) \end{cases} \\ &= \begin{cases} 1, & \text{if } \ x = 1,\ 2 \\ 0, & \text{if } \ x = 3, 4 \end{cases} \end{aligned}$$

Then $a = E_0\phi(X) = .24$ and $b = E_1(1 - \phi(X)) = ..46$, and hence $(a+b)_{\min} = .7$. Note that $\int p_0 \wedge p_1 d\mu = .7$, so this result agrees with the solution of problem 4 of problem set # 4. Also note that for the NP test with $\alpha = .2$, the sum of the two types of error is $.2 + .54 = .74 > .7$.

3. (Problem 3.6, Lehmann and Romano, TSH, page 93.) Suppose that P_0, P_1, and P_2 are the probability distributions assigning to the integers $1, \ldots, 6$ the following probabilities:

x	1	2	3	4	5	6
$p_0(x)$	.03	.02	.02	.01	0	.92
$p_1(x)$	.06	.05	.08	.02	.01	.78
$p_2(x)$	.09	.05	.12	0	.02	.72

Determine whether there exists a level-α test of $H: \; P = P_0$ which is UMP against the alternatives P_1 and P_2 when:
(i) $\alpha = .01$; (ii) $\alpha = .05$; (iii) $\alpha = .07$.

Solution: Here the table of likelihood ratios is as follows:

x	1	2	3	4	5	6
$p_1(x)/p_0(x)$	2	5/2	4	2	∞	78/98
$p_2(x)/p_0(x)$	3	5/2	6	0	∞	72/98

(i) For $\alpha = .01$, the most powerful tests of P_0 versus P_1 and P_2 are of the form

$$\phi_1(x) = 1\{x = 5\} + (1/2)1\{x = 3\},$$
$$\phi_2(x) = 1\{x = 5\} + 1/2)1\{x = 3\},$$

so $\phi_1 = \phi_2$ is Uniformly most powerful.
(ii) For $\alpha = .05$, the most powerful tests of P_0 versus P_1 and P_2 are of the form

$$\phi_1(x) = 1_{\{2,3,5\}}(x) + \gamma(x)1_{\{1,4\}}(x),$$
$$\phi_2(x) = 1_{\{1,3,5\}},$$

so there is no UMP test of P_0 versus P_1 and P_2 at this level.
(iii) For $\alpha = .07$, the most powerful tests of P_0 versus P_1 and P_2 are of the form

$$\phi_1(x) = 1_{\{2,3,5\}}(x) + \gamma(x)1_{\{1,4\}}(x),$$
$$\phi_2(x) = 1_{\{1,2,3,5\}},$$

so by taking $\gamma(x) = 1\{x = 1\}$, $\phi_1(x) = \phi_2(x)$, and this test is Uniformly Most Powerful for testing P_0 versus P_1 and P_2.

4. (Problem 3.7, Lehmann and Romano, TSH, page 94.) Suppose that the distribution of X is given by

x	0	1	2	3
$p_\theta(x)$	θ	2θ	$.9 - 2\theta$	$.1 - \theta$

where $0 < \theta < .1$. For testing $H: \theta = .05$ against $\theta > .05$ at level $\alpha = .05$, determine which of the following tests (if any) is UMP:
(i) $\phi(0) = 1$, $\phi(1) = \phi(2) = \phi(3) = 0$;
(ii) $\phi(1) = .5$, $\phi(0) = \phi(2) = \phi(3) = 0$;
(ii) $\phi(3) = 1$, $\phi(0) = \phi(1) = \phi(2) = 0$.

Solution: The likelihood ratios $P_{\theta'}(X = x)/P_\theta(X = x)$

x	0	1	2	3
$P_\theta(X=x)$	θ	2θ	$.9-.2\theta$	$.1-\theta$
$\frac{P_{\theta'}(X=x)}{P_\theta(X=x)}$	$\frac{\theta'}{\theta}$	$\frac{\theta'}{\theta}$	$\frac{9-20\theta'}{9-20\theta}$	$\frac{1-10\theta}{1-10\theta'}$

It is easy to check that

$$\frac{\theta'}{\theta} > \frac{9-20\theta'}{9-20\theta} > \frac{1-10\theta'}{1-10\theta}$$

Hence this family has monotone decreasing likelihood ratio in x (though not strictly), and strictly decreasing likelihood ratio in

$$\begin{aligned} T(x) &= 1\{x=0\} + 1\{x=1\} + 2\cdot 1\{x=2\} + 3\cdot 1\{x=3\} \\ &= x1\{x>0\} + 1\{x=0\}. \end{aligned}$$

It follows from the Karlin - Rubin theorem that a UMP test of $H : \theta \le \theta_0 = .05$ (of its level) is given by

$$\phi(X) = 1_{[T(X)<k]} + \gamma(X)1_{[T(X)=k]}. \tag{1}$$

(i) Note that the test $\phi_1(X) = 1\{X=0\}$ is of the form (1) with $k=1$ and $\gamma(X) = 1\{X=0\}$ and it has level $\alpha = .05$; hence it is a UMP test of H versus K. The power of ϕ_1 is given by $\beta_1(\theta) \equiv E_\theta\phi_1(X) = \theta$.
(ii) The test $\phi_2(X) = .51\{X=1\}$ is also of the form (1) with $k=1$ and $\gamma(X) = .5\cdot 1\{X=1\}$ and it has level $\alpha = .05$. Hence it is also a UMP test of H versus K. The power of ϕ_2 is given by $\beta_2(\theta) \equiv E_\theta\phi_2(X) = \theta$.
(iii) The test $\phi_3(X) = 1\{X=3\}$ is clearly not of the form (1). It has power function $\beta_3(\theta) = E_\theta\phi_3(X) = .1-\theta$, so $\beta_3(.05) = .05$, but $\beta_3(\theta) > .05$ for $\theta < .05$ while $\beta_3(\theta) < .05$ for $\theta > \theta_0 = .05$. In fact, this is a UMP test of $\tilde{H} : \theta \ge \theta_0$ versus $\tilde{K} : \theta < \theta_0$.

5. Suppose that $X_1, \ldots, X_n$ are i.i.d. $N(\theta, \sigma^2)$.
(a) Suppose that $\sigma = \sigma_0$ is known. Consider testing $H : \theta = \theta_0 = 0$ versus $K : \theta = \theta_1 = 1$. In the spirit of chapter 5, plot $(R(\theta_0, \phi), R(\theta_1, \phi))$ for your favorite family of tests ϕ. Find the entire risk body and plot it.
(b) What happens to the risk body as n grows or as $\sigma_0 \to 0$?
(c) What happens to the risk body as θ_1 decreases toward $\theta_0 = 0$?
(d) What happens to the risk bodies $\{(R(\theta_0, \phi), R(\theta_{1,n}, \phi)) : \ n \ge 1\}$ when $\theta_1 \equiv \theta_{1,n} \equiv \theta_0 + cn^{-1/2}$?

Solution: (a) If $X_1, \ldots, X_n$ are i.i.d. $N(\theta, \sigma_0)$, to find optimal tests ϕ we can reduce (by sufficiency) to consideration of $\overline{X} \sim N(\theta, \sigma_0^2/n)$. My favorite family of tests (in fact the most powerful tests) of H versus K are the tests $\phi_c(\underline{X}) = 1\{\overline{X} > c\}$.For these tests

$$\begin{aligned} R(0, \phi_c) &= E_0\phi_c(\underline{X}) = P_0(\overline{X} > c) \\ &= P_0(\sqrt{n}(\overline{X} - 0)/\sigma_0 > \sqrt{n}c/\sigma_0) \\ &= 1 - \Phi(\sqrt{n}c/\sigma_0) \end{aligned}$$

and

$$\begin{aligned} R(1, \phi_c) &= E_1(1 - \phi_c(\underline{X})) \\ &= P_1(\overline{X} \le c) = P_1(\sqrt{n}(\overline{X} - 1) \le \sqrt{n}(c - 1)) \\ &= \Phi(\sqrt{n}(c - 1)/\sigma_0). \end{aligned}$$

Since these tests are MP for testing H versus K, there are no points with risks below the curve given by $\{(R(0, \phi_c), R(1, \phi_c)) : c \in \mathbb{R}\}$; this is the lower boundary of the risk body. Note that the tests $\phi_{\text{ignore}}(\underline{X}) \equiv \alpha$ have risks $R(0, \phi_{\text{ignore}}) = \alpha$, $R(1, \phi_{\text{ignore}}) = 1 - \alpha$. Thus the line $\{(\alpha, 1 - \alpha) : \alpha \in [0, 1]\}$ is in the risk body. Furthermore, note that the tests $\phi'_c(\underline{X}) \equiv 1 - \phi_c(\underline{X}) = 1\{\overline{X} \le c\}$ are MP for testing $H : \theta = 0$ versus $K' : \theta = \theta_1 < 0$, and by the Karlin - Rubin theorem these tests minimize the power function at points $\theta = \theta_1$ in the class of all tests with fixed power function (say at α) at $\theta = \theta_0$. Since

$$\text{Power}_{\phi'_c}(\theta) = E_\theta \phi'_c = 1 - R(\theta, \phi_c),$$

this says that the tests ϕ'_c maximize $R(1, \phi_c)$ over tests ϕ with $R(0, \phi) = \alpha$. Hence there are no points in the risk body with risks above the curve given by $\{(1 - R(0, \phi_c), 1 - R(1, \phi_c)) : c \in \mathbb{R}\}$.
(b) As n grows or $\sigma_0 \to 0$ the risk body expands out toward the boundary of the square $[0, 1]^2$; see the plots below.
(c) As $\theta_1 \to \theta_0 = 0$, the risk body contracts toward the diagonal line $(\alpha, 1 - \alpha)$ – since the testing problem becomes harder. See the plots below.

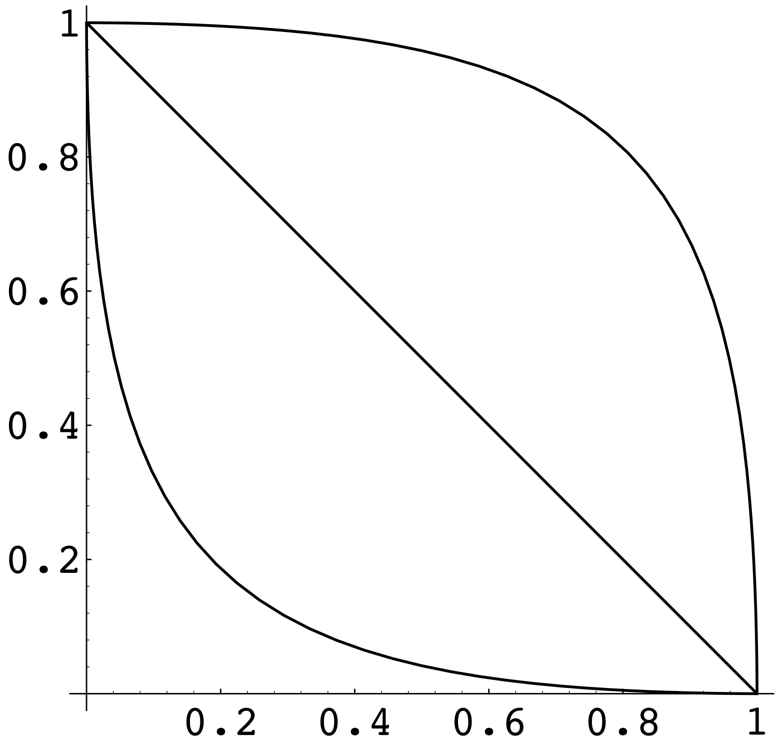

Figure 1: Risks for normal mean test, $n = 3$, $\sigma_0 = 1$, $\theta_1 = 1$

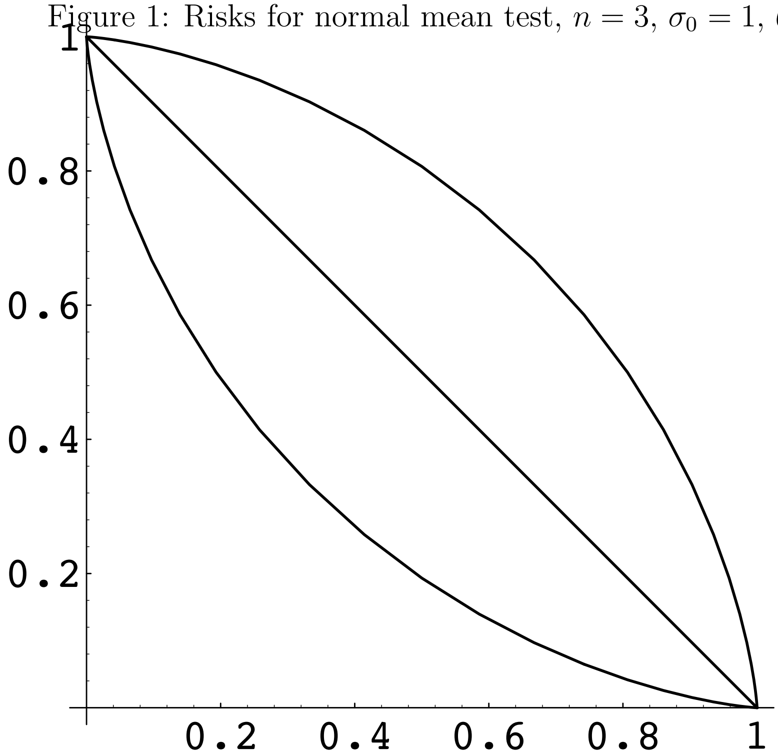

Figure 2: Risks for normal mean test, $n = 3$, $\sigma_0 = 1$, $\theta_1 = .5$

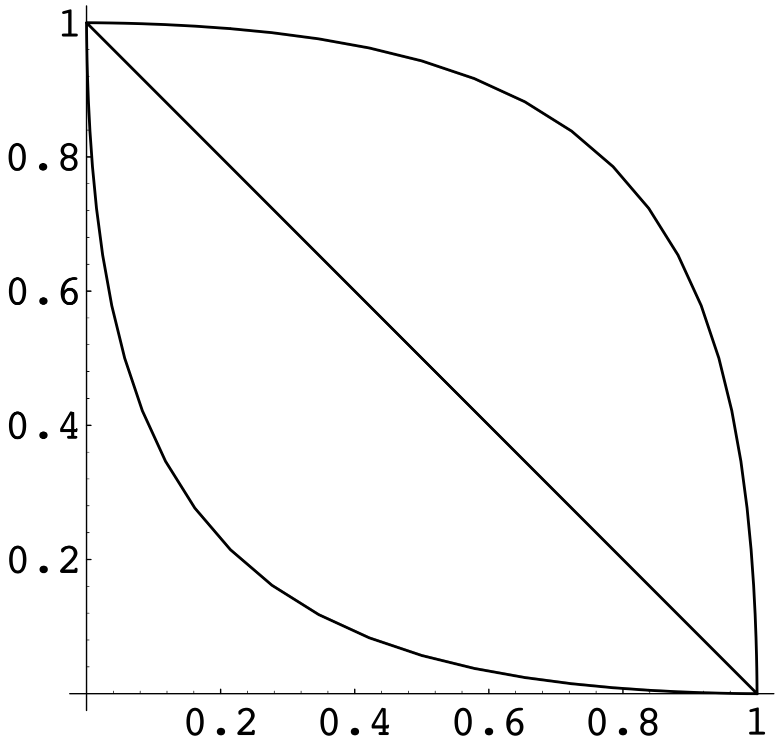

Figure 3: Risks for normal mean test, $n = 10$, $\sigma_0 = 1$, $\theta_1 = .5$

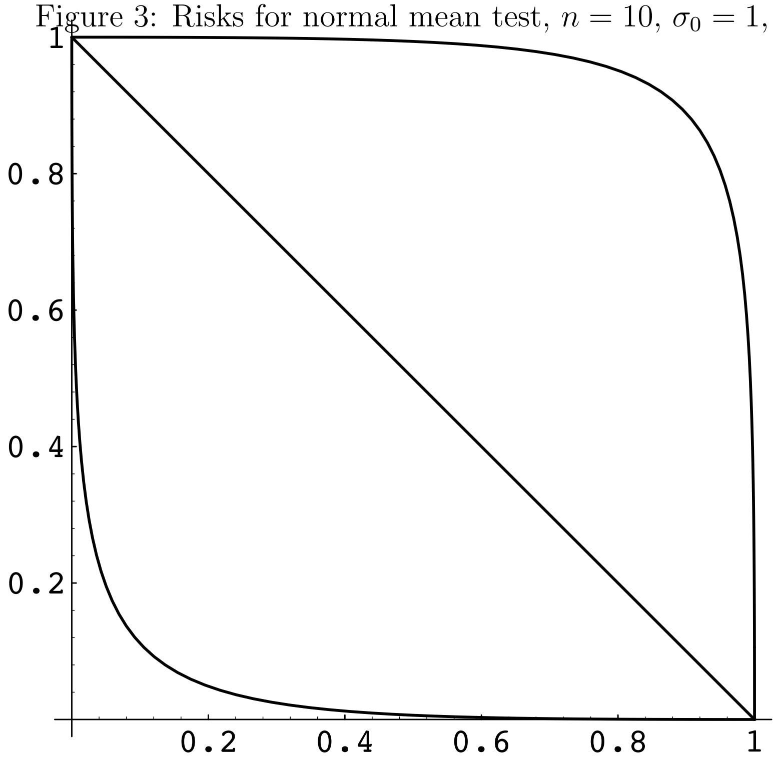

Figure 4: Risks for normal mean test, $n = 25$, $\sigma_0 = 1$, $\theta_1 = .5$