Statistics 582, Problem Set 8 Solutions

Wellner; 2/28/2018

1. Consider the Locally Most Powerful test ϕ for testing $H : \theta \le 0 \equiv \theta_0$ versus $K : \theta > 0 = \theta_0$ in Example 6.1.5.
(a) Suggest two different approximations to the power of this test, one for local alternatives (of the form $\theta_n = t/\sqrt{n}$ with $t > 0$), and the other for fixed alternatives, $\theta > 0$.
(b) What is the behavior of each of these two approximations for large values of θ? Which of them shows that the power function decreases to 0 as $\theta \to \infty$? Why?

Solution: (a) The test is "reject H if $\sqrt{n}\,\overline{Y}_n > 2^{-1/2} z_\alpha$" where $Y_i \equiv 2X_i/(1+X_i^2)$ are i.i.d. and $X_i \sim$ Cauchy $(\theta, 1)$. Thus under P_θ, by using contour integration and Cauchy's formula, or by using Mathematica, Maple, or your favorite symbolic manipulation program,

$$\begin{aligned} m(\theta) \equiv E_\theta Y_i &= \int_{-\infty}^{\infty} \frac{2x}{1+x^2} p_\theta(x)dx = \int_{-\infty}^{\infty} \frac{2x}{1+x^2}\frac{1}{\pi}\frac{1}{1+(x-\theta)^2}dx \\ &= \frac{2\theta}{4+\theta^2}, \end{aligned}$$

and

$$\begin{aligned} \sigma^2(\theta) &\equiv Var_\theta(Y_i) = E_\theta Y_i^2 - m^2(\theta) = \frac{2(4+3\theta^2)}{(4+\theta^2)^2} - \left(\frac{2\theta}{4+\theta^2}\right)^2 \\ &= \frac{2}{4+\theta^2}\,. \end{aligned}$$

For local alternatives $\theta = \theta_n = t/\sqrt{n}$, we have

$$\begin{aligned} \text{Power}(\theta_n) &= P_{\theta_n}(\sqrt{n}\,\overline{Y}_n > 2^{-1/2} z_\alpha) \\ &= P_{\theta_n}(\sqrt{n}(\overline{Y}_n - m(\theta_n)) \ge 2^{-1/2} z_\alpha - \sqrt{n}(m(\theta_n) - m(0))) \\ &\to P(2^{-1/2} Z \ge 2^{-1/2} z_\alpha - m'(0)t) \end{aligned}$$

where

$$\begin{aligned} m'(0) &= \int_{-\infty}^{\infty} \frac{2x}{1+x^2}\frac{d}{d\theta} p_\theta(x)|_{\theta=0} dx \\ &= \int_{-\infty}^{\infty} \dot{l}_\theta(x;0)\dot{l}_\theta(x;0) p_\theta(x;0)dx = I(\theta) = 1/2\,. \end{aligned}$$

Hence we have

$$\text{Power}(\theta_n) \to P(Z > z_\alpha - 2^{-1/2}t) = 1 - \Phi(z_\alpha - 2^{-1/2}t)\,.$$

This approximation to the power function increases monotonically from α at $t = 0$ to 1 at $t = \infty$ (effectively when $t > 2^{1/2} \cdot 4$). Note that this result is very much in qualitative agreement with corollary 4.2.4 from Statistics 581.

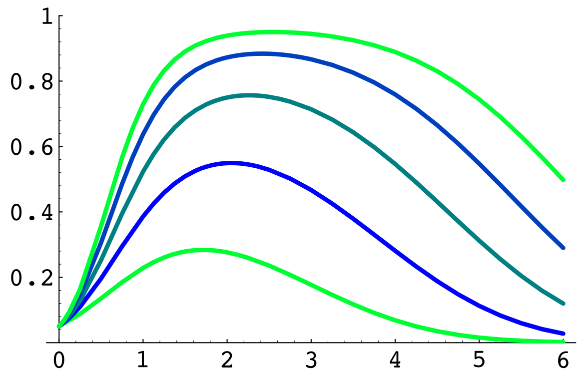

Figure 1: Plots of fixed θ power approximations for $n = 3, 6, 9, 12, 15$.

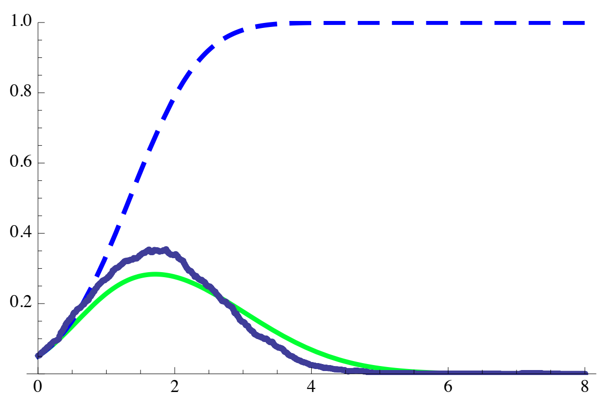

Figure 2: Local (dashed) and fixed (green) θ power approximations, $n = 3$; Monte Carlo estimate of true power (blue)

(b) For fixed alternatives $\theta > 0$ we have

$$\begin{aligned}
\text{Power}(\theta) &= P_\theta(\sqrt{n}\,\overline{Y} > 2^{-1/2}z_\alpha) \\
&= P_\theta(\sqrt{n}(\overline{Y}_n - m(\theta)) > 2^{-1/2}z_\alpha - \sqrt{n}m(\theta)) \\
&\doteq P(Z > (2^{-1/2}z_\alpha - \sqrt{n}m(\theta))/\sigma(\theta)) \\
&= 1 - \Phi((2^{-1/2}z_\alpha - \sqrt{n}m(\theta))/\sigma(\theta))\,.
\end{aligned}$$

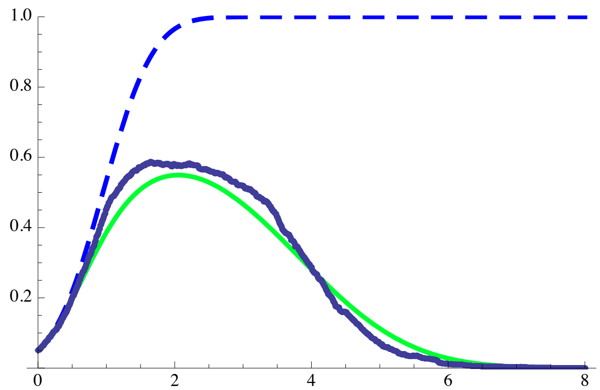

Figure 3: Local (dashed) and fixed (green) θ power approximations, $n = 6$; Monte Carlo estimate of true power (blue)

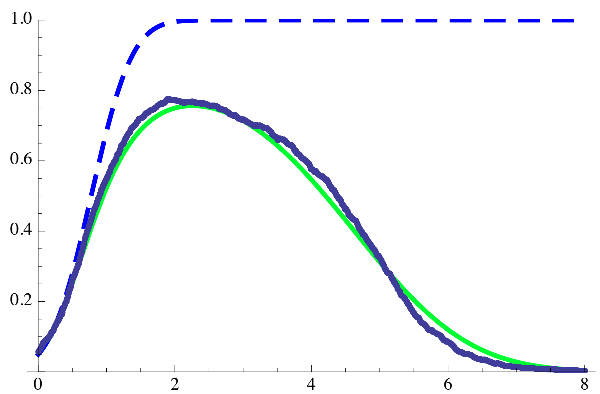

Figure 4: Local (dashed) and fixed (green) θ power approximations, $n = 9$; Monte Carlo estimate of true power (blue)

where

$$\frac{m(\theta)}{\sigma(\theta)} = \frac{\sqrt{2\theta^2}}{\sqrt{4+\theta^2}}$$

which increases from 0 at $\theta = 0$ to a maximum of $1/\sqrt{2}$ at $\theta = 2$, and then tends to $\sqrt{2}$ as θ increases to ∞. Thus our fixed alternative approximation to the power function is completely determined by the function $(2^{-1/2}z_\alpha - \sqrt{n}m(\theta))/\sigma(\theta)$, and this first decreases as θ increases (as it should if the power is to increase), but then it reaches a minimum and increases thereafter (approximately as $2^{-1}z_\alpha\theta - \sqrt{2n}$ as $\theta \to \infty$). Hence this approximation to the power decreases to zero just as we argued that it must in class. See Figures 1-6 which give plots of the two approximations for $\alpha = .05$ and $n = 3, 6, 9, 12, 15$, together with Monte-Carlo estimates of the true power function based on $m = 1000$ Monte-Carlo replica-

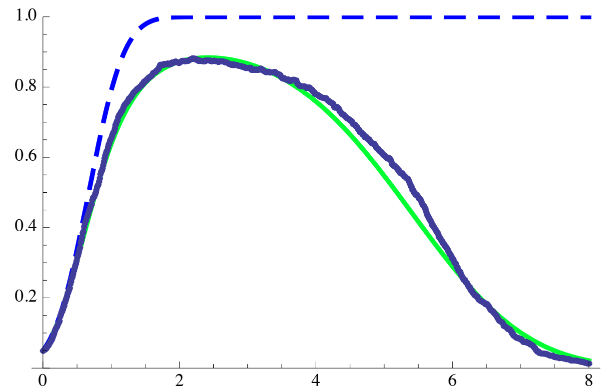

Figure 5: Local (dashed) and fixed (green) θ power approximations, $n = 12$; Monte Carlo estimate of true power (blue)

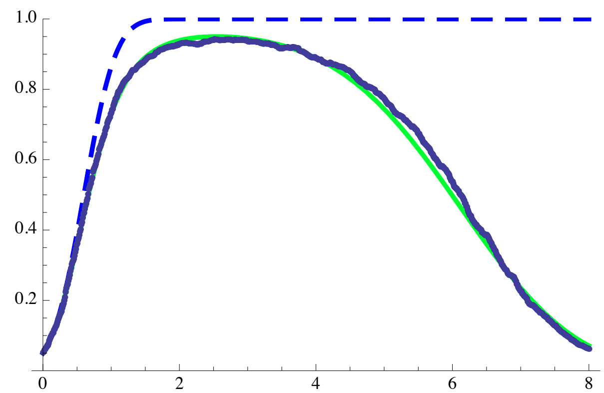

Figure 6: Local (dashed) and fixed (green) θ power approximations, $n = 15$; Monte Carlo estimate of true power (blue)

tions at each sample size (dark blue). Note that the two approximations agree for θ's close to 0, but the local approximation is always monotone increasing, while the approximations with θ fixed show the approximate power decreasing to 0 as $\theta \to \infty$ as we know it must. Also note that the fixed alternative approximation to the power function based on the CLT is reasonably accurate for $n \geq 9$.

2. This is a continuation of the previous problem.
(a) Suppose you decide to base a test of the hypotheses in the setting of Example 6.1.5 on the sample mean $\overline{X}_n$. Show how to carry out this test at exact size $0 < \alpha < 1/2$, and compute its power function $\beta_\phi(\theta)$. [Hint: Recall that for $\theta_0 = 0$, and Cauchy data, $\overline{X}_n \stackrel{d}{=} X_1$.]
(b) Alternatively, consider a test of the same hypotheses on the sample median

$\mathbb{F}_n^{-1}(1/2)$, or perhaps more simply, the proportion of observations exceeding $\theta_0 = 0$. Show how to carry out this test at exact and approximate size $\alpha \in (0, 1/2)$. Compute the power function of your new test and show that it has a monotone non-decreasing power function $\beta_\phi(\theta)$.
(c) Compare the power functions (or best approximations thereof) for the tests in (a) and (b) and problem 1.

Solution: (a) Consider $\overline{X}_n \equiv n^{-1} \sum_{i=1}^n X_i$. Then, since $X_i \stackrel{d}{=} V_i + \theta$ where V_i are i.i.d. standard Cauchy random variables with $Ee^{itV_1} = \exp(-|t|)$ (see e.g. Shorack, *Probability for Statisticians*, page 343), it follows that

$$\begin{aligned}
\phi_{\overline{X}_n}(t) &= E\exp(itn^{-1}\sum_1^n X_i) = \phi_{X_1}(t/n)^n = \big(E\exp(itn^{-1}X_1)\big)^n \\
&= \big(E\exp(itn^{-1}(V_1+\theta))\big)^n = \phi_{V_1}(t/n)^n \cdot e^{it\theta} = \\
&= \exp(-|t|)e^{it\theta} = \phi_{X_1}(t);
\end{aligned}$$

i.e. $\overline{X}_n \stackrel{d}{=} X_1$. Thus the test $\phi(\underline{X}) = 1\{\overline{X}_n > k\}$ has

$$\alpha = E_0\phi(\underline{X}) = P_0(\overline{X}_n > k = P_0(X_1 > k)$$

if $k = \tan((1/2 - \alpha)\pi) \equiv k_\alpha$, and then

$$\begin{aligned}
\beta_\phi(\theta) &= P_\theta(\overline{X}_n > k_\alpha) = P_\theta(X_1 - \theta > k_\alpha - \theta) \\
&= P_0(V_1 > k_\alpha - \theta) = \int_{k_\alpha-\theta}^\infty \frac{1}{\pi}\frac{1}{1+y^2}dy
\end{aligned}$$

which is monotone increasing as a function of θ; see Figure 7. On the other hand, this test does not yield increasing power as the sample size n increases: it has the same power function for all n!
(b) Let $Z_i \equiv 1\{X_i > 0\}$ for $i = 1, \ldots, n$. Thus the Z_i's are independent Bernoulli $p = p_\theta$ random variables with

$$\begin{aligned}
p_\theta &= P_\theta(X_i > 0) = P_\theta(X_i - \theta > -\theta) = P_0(Y_i > -\theta) = \int_{-\theta}^\infty \frac{1}{\pi}\frac{1}{1+y^2}dy \\
&\nearrow 1 \text{ as } \theta \to \infty.
\end{aligned}$$

Note that testing $H : \theta \le \theta_0 = 0$ versus $K : \theta > 0$ is equivalent to testing $H' : p \le p_0 \equiv 1/2$ versus $K' : p > p_0 = 1/2$. Now

$$T_n \equiv \sum_1^n Z_i = n(1 - \mathbb{F}_n(0)) \sim \text{Binomial}(n, p)$$

is a one-parameter exponential family with density function $p(\cdot;p)$ with respect to counting measure on $\mathbb{N}$ given by

$$\begin{aligned} p(y;p) &= P_p(T_n = y) = \binom{n}{y} p^y (1-p)^{n-y} = \binom{n}{y} \left(\frac{p}{1-p}\right)^y (1-p)^n \\ &= c(p)\exp(Q(p)y)h(y) \end{aligned}$$

with $Q(p) \equiv \log(p/(1-p))$, $c(p) \equiv (1-p)^n$, and $h(y) = \binom{n}{y}$, so it has monotone likelihood ratio in y. Thus by the Karlin-Rubin theorem, the test

$$\varphi(\underline{X}) = \phi(\underline{Z}) = \begin{cases} 1 & \text{if } T_n = \sum_1^n Z_i > k \\ \gamma & \text{if } T_n = k \\ 0 & \text{if } T_n < k, \end{cases}$$

where k and γ are chosen so that $E_{p_0=1/2}\phi(\underline{Z}) = \alpha$, has power function

$$\beta_\phi(p_\theta) \equiv \tilde{\beta}_\varphi(\theta)$$

which is monotone increasing in p_θ and θ. This test improves with increasing sample size n; see Figure 7

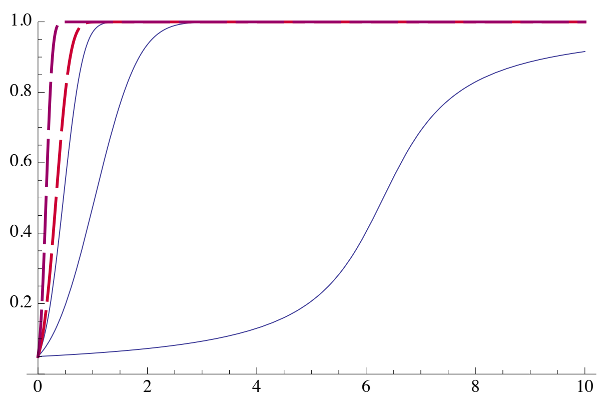

Figure 7: Plot comparing local power approximations for the Binomial test (solid), locally most powerful test (dashed), and the mean test, for $n = 10,\ 50$.

3. Let X and Y be random variables with joint density

$$p_{X,Y}(x,y) = \lambda\mu\exp(-\lambda x - \mu y)1_{(0,\infty)}(x)1_{(0,\infty)}(y).$$

(a) Find a UMP unbiased test of size $\alpha = .2$ for testing $H_0 : \lambda \le \mu + 1$ versus $H_1 : \lambda > \mu + 1$.
(b) Find a UMP unbiased test of size $\alpha = .2$ for testing $H_0 : \lambda = \mu$ versus $H_1 : \lambda \ne \mu$.

(c) Find a UMP unbiased test of size $\alpha = .2$ for testing $H_0 : \lambda \geq 2\mu$ versus $H_1 : \lambda < 2\mu$.
(d) What happens when $X_1, \ldots, X_m$ are i.i.d. Exponential(λ) and $Y_1, \ldots, Y_n$ are i.i.d. Exponential(μ)?

Solution: When $X \sim \exp(\lambda)$ and $Y \sim \exp(\mu)$ we have

$$p_{\lambda,\mu}(x,y) = \lambda\mu \exp(-\lambda x - \mu y) 1_{(0,\infty)}(x) 1_{(0,\infty)}(y) \,.$$

(a) For testing $H : \lambda \leq \mu + 1$ versus $K : \lambda > \mu + 1$ we rewrite the density as follows:

$$\begin{aligned} p_{\lambda,\mu}(x,y) &= \lambda\mu \exp(-\lambda x - \mu y) \\ &= \lambda\mu \exp((\lambda - \mu)y - \lambda(x+y)) \\ &= \lambda\mu \exp(\theta U(x,y) + \xi T(x,y)) \end{aligned}$$

where $\theta \equiv \lambda - \mu$, $U(x,y) \equiv y$, $\xi \equiv -\lambda$, and $T(x,y) \equiv x + y$. Since $\lambda \leq \mu + 1$ is equivalent to $\lambda - \mu = \theta \leq 1 \equiv \theta_0$, our theory for exponential families applies, and the UMP unbiased test of H versus K is given by

$$\phi(X,Y) = \begin{cases} 1 & \text{if } \; Y > c_\alpha(T) \\ \gamma(T) & \text{if } \; Y = c_\alpha(T) \\ 0 & \text{if } \; Y < c_\alpha(T) \end{cases}$$

where c_α and $\gamma(\alpha)$ satisfy $E\{\phi(X,Y)|T\} = \alpha$. In this case, the conditional distribution of Y given $T = X + Y$ on the boundary $\Theta_B = \{(\lambda - 1, \lambda) : \lambda \geq 1\}$ is given by

$$f_{Y|T}(y|t) = \frac{e^y}{e^t - 1} 1_{[0,t]}(y) \,.$$

Therefore $1 - F_{Y|T}(y|t) = 1 - (e^y - 1)/(e^t - 1)$ and for $\alpha = .2$ the critical point for the conditional test is given by

$$c_\alpha(T) = \log(\exp(T) - (\exp(T) - 1)/5) = \log((4/5)\exp(T) + 1/5) \,, \qquad \gamma(T) = 0 \,.$$

(b) For testing $H : \lambda = \mu$ versus $K : \lambda \neq \mu$, the same rewrite of the density as in (a) works. Now we have $\lambda = \mu$ is equivalent to $\mu - \lambda = 0 \equiv \theta_0$, and $\lambda \neq \mu$ is equivalent to $\mu - \lambda \neq 0 \equiv \theta_0$, so our theory for exponential families applies, and the UMP unbiased test of H versus K is given by

$$\phi(X,Y) = \begin{cases} 1 & \text{if } \; Y > c_2(T) \; \text{ or } \; Y < c_1(T) \\ \gamma_i(T) & \text{if } \; Y = c_1(T) \; \text{ or } \; Y = c_2(T) \\ 0 & \text{if } \; c_1(T) < Y < c_2(T) \end{cases}$$

where the c_1, c_2, γ_1 and γ_2 are determined so that $E\{\phi(X,Y)|T\} = \alpha$. In this case the conditional distribution of Y given T on $\Theta_B = \{(\lambda,\lambda) : \lambda \geq 0\}$ is Uniform$(0,T)$, and hence the conditional distribution of Y/T given T is Uniform$(0,1)$, and this is independent of T. Hence the UMPU test of H versus K of size .2 is given by "reject H if $Y/T < .1$ or $Y/T > .9$".
(c) For testing $H : \lambda \geq 2\mu$ versus $K : \lambda < 2\mu$, we need a somewhat different rewrite of the joint density. Now

$$\begin{aligned} p_{\lambda,\mu}(x,y) &= \lambda\mu \exp(-\lambda x - \mu y) \\ &= \lambda\mu \exp(-(\lambda - 2\mu)x - \mu(2x + y)) \\ &= \lambda\mu \exp(\theta U(x,y) + \xi T(x,y)) \end{aligned}$$

where $\theta \equiv 2\mu - \lambda$, $U(x,y) \equiv x$, $\xi \equiv -\mu$, and $T(x,y) \equiv 2x + y$. Since $\lambda \geq 2\mu$ is equivalent to $2\mu - \lambda \equiv \theta \leq 0 \equiv \theta_0$, (and $\lambda < 2\mu$ is equivalent to $2\mu - \lambda = \theta > 0 \equiv \theta_0$), our theory for exponential families applies, and the UMP unbiased test of H versus K is given by

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

where $c_(T)$ and $\gamma(T)$ satisfy $E\{\phi(X,Y)|T\} = \alpha$. In this case the conditional distribution of X given T is Uniform$(0,T/2)$, so $2X/T$ is Uniform$(0,1)$ and independent of T. Hence the UMPU test of H versus K of size $\alpha = .2$ is given by "reject H if $2X/T > .8$".

When we observe $X_1,\ldots,X_m$ are i.i.d. Exponential(λ) and $Y_1,\ldots,Y_n$ are i.i.d. Exponential(μ), then the distribution of the observations is given by

$$\begin{aligned} p_{\lambda,\mu}(\underline{x},\underline{y}) &= \lambda^m\mu^n \exp(-\lambda\sum_{i=1}^m x_i - \mu\sum_{j=1}^n y_j) \\ &= \lambda^m\mu^n \exp((\lambda-\mu)\sum_{j=1}^n y_j - \lambda(\sum_{i=1}^m x_i + \sum_{j=1}^n y_j)) \\ &= \lambda^m\mu^n \exp(\theta U(\underline{x},\underline{y}) + \xi T(\underline{x},\underline{y})) \end{aligned}$$

where $\theta \equiv \lambda - \mu$, $U(\underline{x},\underline{y}) \equiv \sum y_j$, $\xi \equiv -\lambda$, and $T(\underline{x},\underline{y}) \equiv \sum x_i + \sum y_j)$. This rewrite works for (a) and (b), and a similar rewrite works for (c) with the new $U = \sum x_i$, $T = 2\sum x_i + \sum y_j$. The form of the tests in (a) - (c) remains the same with the new U and T, and all that remains is to calculate the conditional distributions of U given T. In (a) this density is given by

$$f_{U|T}(u|t) = \frac{u^{n-1}(t-u)^{m-1}e^u}{\int_0^t v^{n-1}(t-v)^{m-1}e^v dv}\,.$$

In (b) it is easily found that $V \equiv U/T \sim \text{Beta}(n, m)$, and the test can be carried out unconditionally using tables of the Beta distributions. In (c) $2U \equiv \sum 2X_i \sim \text{Gamma}(m, \mu)$ and $V \equiv \sum Y_j \sim \text{Gamma}(n, \mu)$ on the boundary $\lambda = 2\mu$, so $2\mu U \sim \chi^2_{2m}$ and $\mu V \sim \chi^2_{2n}$. Therefore

$$\frac{2U/(2m)}{V/(2n)} = \frac{n}{m} 2\frac{U}{V} \sim F_{2m,2n}$$

and since $2U/T = 2U/(2U + V) = (2U/V)(1 + (2U/V))$ is a monotone increasing function of $2U/V$, the UMPU test can be carried out unconditionally using tables of the $F_{2m,2n}$ distributions.

4. Lehmann and Romano, TSH, problem 4.3, page 139, modified: Let $X \sim \text{Binomial}(n, p)$, and consider testing $H : p = p_0$ versus $K : p \neq p_0$ at level $\alpha = \alpha$. Determine the boundary values of the UMP unbiased test for $n = 20$ with $\alpha = .05$, $p_0 = .2$, and with $\alpha = .10$, $p_0 = .4$. In each case plot the power functions of both the unbiased and the equal-tails test.

   Solution: For the first scenario: $n = 20$, $\alpha = .05$, $p_0 = .2$, we find, using the calculations in Lehmann and Romano, TSH, page 113, the UMP unbiased test is $\phi(x) = \gamma_1 1\{x = 1\} + \gamma_2 1\{x = 8\} + 1\{x \geq 9\}$ +1{x=0}with $c_1 = 1$, $c_2 = 8$, $\gamma_1 = .0.2999$, and $\gamma_2 = .5054$. The corresponding equal tails test has $c_1 = 1$, $c_2 = 8$, $\gamma_1 = 0.2337$, and $\gamma_2 = 0.6777$. The power functions of these two tests are shown in Figures 8 and 9.

   For the second scenario: $n = 20$, $\alpha = .10$, $p_0 = .4$, we find, using the calculations in Lehmann and Romano, TSH, page 113, the UMP unbiased test is $\phi(x) = \gamma_1 1\{x = 5\} + \gamma_2 1\{x = 12\} + 1\{x \geq 13\} + 1\{x \leq 4\}$ with $c_1 = 5$, $c_2 = 12$, $\gamma_1 = 0.0033$, and $\gamma_2 = 0.782377$.

   The corresponding equal tails test has $c_1 = 4$, $c_2 = 12$, $\gamma_1 = 0.9738$, and $\gamma_2 = 0.8161$. The power functions of these two tests are shown in Figure 10; since the binomial distribution is quite close to being symmetric for $p = .4$, they do not differ substantially.

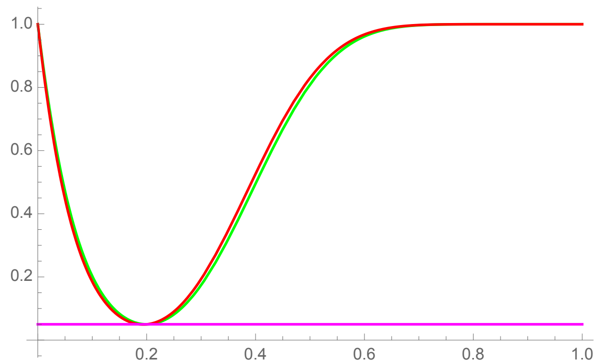

Figure 8: Power functions, UMP unbiased test (green) and equal tails test (red): $n = 20$, $p_0 = .2$, $\alpha = .05$

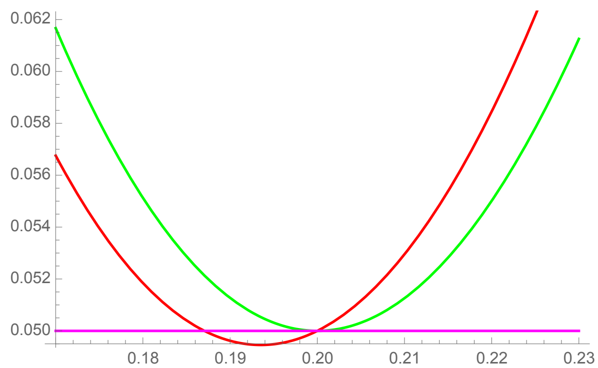

Figure 9: Power functions, UMP unbiased test (green) and equal tails test (red): $n = 20$, $p_0 = .2$, $\alpha = .05$

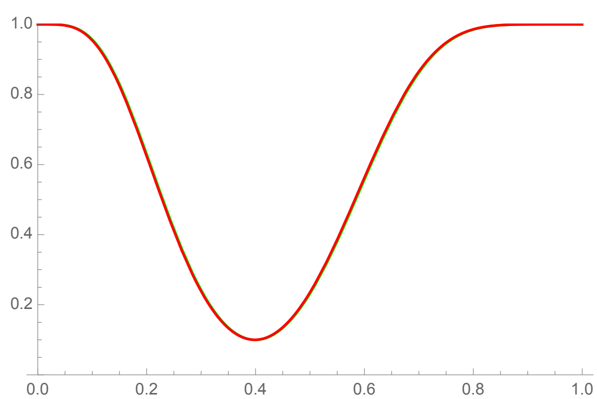

Figure 10: Power functions, UMP unbiased test (green) and equal tails test (red): $n = 20$, $p_0 = .4$, $\alpha = .10$