ECE531 Homework Assignment Number 5

Solution

Make sure your reasoning and work are clear to receive full credit for each problem.

1. 4 points. Poor textbook Chapter II, Problem 13.

   Solution: Since this is a binary composite HT problem, we want to find a Bayes decision rule

$$\delta^{B\pi}(y) = \arg\min_{i \in \{0,1\}} g_i(y, \pi)$$

   where

$$\begin{aligned} g_0(y,\pi) &= \int_\beta^\infty \alpha e^{-\alpha\theta} \theta e^{-\theta y} d\theta = \frac{\alpha(\alpha\beta + y\beta + 1)e^{-\beta(\alpha+y)}}{(\alpha+y)^2} \\ g_1(y,\pi) &= \int_0^\beta \alpha e^{-\alpha\theta} \theta e^{-\theta y} d\theta = \frac{\alpha\left[1 - (\alpha\beta + y\beta + 1)e^{-\beta(\alpha+y)}\right]}{(\alpha+y)^2}. \end{aligned}$$

   Note the common term $\frac{\alpha}{(\alpha+y)^2}$. Denote $f(y) = (\alpha\beta + y\beta + 1)e^{-\beta(\alpha+y)}$ where $\alpha > 0$ and $\beta > 0$ are both fixed and known. Thus we have the Bayes decision rule

$$\delta^{\pi B}(y) = \begin{cases} 1 & \text{if } f(y) \geq \frac{1}{2} \\ 0 & \text{if } f(y) < \frac{1}{2} \end{cases}$$

2. 6 points. Consider n uniform i.i.d. $Y_i \sim \mathcal{U}(0, x)$ random variables with $x > 0$ and hypotheses

$$\begin{aligned} \mathcal{H}_0 &: x \leq \lambda \\ \mathcal{H}_1 &: x > \lambda \end{aligned}$$

   (a) Define $T(Y) = \max_i Y_i$ and find the significance level of the following decision rule

$$\rho(Y) = \begin{cases} 1 & \text{if } T(Y) > \lambda \\ \alpha & \text{if } T(Y) \leq \lambda \end{cases}$$

   where $\rho(Y) = \alpha$ means that we decide $\mathcal{H}_1$ with probability α. Explain the intuition behind ρ. Derive the power function of ρ. Is ρ a UMP decision rule of size α?

   Solution: The significance level of the decision rule can be written as

$$\begin{aligned} P_{\mathbf{fp}} &= \text{Prob}[\max_i Y_i > \lambda \,|\, x \leq \lambda] + \alpha \text{Prob}[\max_i Y_i \leq \lambda \,|\, x \leq \lambda] \\ &= 0 + \alpha \cdot 1 \\ &= \alpha \end{aligned}$$

When $T(Y) > \lambda$ we must have $x > \lambda$. Hence $\mathcal{H}_0$ can not be true when $T(Y) > \lambda$. When $T(Y) \leq \lambda$, it is possible for either hypothesis to be true. By choosing $\mathcal{H}_1$ with probability α, we achieve the desired false positive probability (significance level) of α.

Given n i.i.d. observations, the power function for $x > \lambda$ can be written as

$$\begin{aligned}
\beta(x) &= \text{Prob}[\max_i Y_i > \lambda \,|\, x > \lambda] + \alpha \text{Prob}[\max_i Y_i \leq \lambda \,|\, x > \lambda] \\
&= \text{Prob}[\max_i Y_i > \lambda \,|\, x > \lambda] + \text{Prob}[\max_i Y_i \leq \lambda \,|\, x > \lambda] - (1-\alpha)\text{Prob}[\max_i Y_i \leq \lambda \,|\, x > \lambda] \\
&= 1 - (1-\alpha)\text{Prob}[\max_i Y_i \leq \lambda \,|\, x > \lambda] \\
&= 1 - (1-\alpha)\left(\frac{\lambda}{x}\right)^n.
\end{aligned}$$

Hence the overall power function is

$$\beta(x) = \begin{cases} 1 - (1-\alpha)\left(\frac{\lambda}{x}\right)^n & \text{if } x > \lambda \\ \alpha & \text{if } x \leq \lambda \end{cases}$$

We will confirm that this is a UMP decision rule in part (c).

(b) Show that the vector observation $Y = (Y_0, \ldots, Y_{n-1})$ has a monotone likelihood ratio in $T(Y)$.

Solution: For any $x_0 \in (0, \infty)$ and $x_1 \in (0, \infty)$ with $x_1 > x_0$, we can compute

$$L_{x_1/x_0}(y) = \frac{p_{x_1}(y_0, \ldots, y_{n-1})}{p_{x_0}(y_0, \ldots, y_{n-1})} = \begin{cases} \frac{(\frac{1}{x_1})^n}{(\frac{1}{x_0})^n} = \left(\frac{x_0}{x_1}\right)^n & \text{if } 0 < T(y) \leq x_0 \\ \infty & \text{if } x_0 < T(y) \leq x_1 \end{cases}$$

This result shows that we have a monotone likelihood ratio in $T(y)$.

(c) Find a UMP decision rule of size α and compare its power function to that of ρ.

Solution: Since we have a monotone likelihood ratio, we know the UMP decision rule must have the form

$$\rho(y) = \begin{cases} 1 & \text{if } T(y) > \tau \\ \gamma & \text{if } T(y) = \tau \\ 0 & \text{if } T(y) < \tau \end{cases}.$$

We just need to solve for τ and γ so that $P_{\mathbf{fp},x=\lambda} = \alpha$. Since $T(Y)$ is a continuous random variable, we know that we don't need to worry about the randomization. We just need to solve for τ.

When $\tau \geq \lambda$, we have $P_{\mathbf{fp},x=\lambda} = 0$. So that isn't very useful.

When $\tau < \lambda$, we have

$$\begin{aligned}
P_{\mathbf{fp},x=\lambda} &= \text{Prob}[\max_i Y_i > \tau \,|\, x = \lambda] \\
&= 1 - \text{Prob}[\max_i Y_i \leq \tau \,|\, x = \lambda] \\
&= 1 - \left(\frac{\tau}{\lambda}\right)^n.
\end{aligned}$$

Setting this equal to α yields the desired result $\tau = (1-\alpha)^{\frac{1}{n}}\lambda$. Thus

$$\rho(y) = \begin{cases} 1 & \text{if } T(y) \geq (1-\alpha)^{\frac{1}{n}}\lambda \\ 0 & \text{if } T(y) < (1-\alpha)^{\frac{1}{n}}\lambda. \end{cases}$$

The power function

$$\beta(x) := \text{Prob}[\rho \text{ decides } 1 \mid \text{state is } x]$$

We need to evaluate this in three regions:

- When $x < (1-\alpha)^{\frac{1}{n}}\lambda$, we always have $T(y) < (1-\alpha)^{\frac{1}{n}}\lambda$, hence we always decide $\mathcal{H}_0$. Hence $\beta(x) = 0$. Note that $\beta(x) = P_{\mathbf{fp},x}$ in this region since $x \in \mathcal{X}_0$.
- When $(1-\alpha)^{\frac{1}{n}}\lambda \le x \le \lambda$, the statistic $T(y)$ will be greater than or equal to the threshold $(1-\alpha)^{\frac{1}{n}}\lambda$ with probability $1-(1-\alpha)(\frac{\lambda}{x})^n$. Hence $\beta(x) = 1-(1-\alpha)(\frac{\lambda}{x})^n$. Again note that $\beta(x) = P_{\mathbf{fp},x}$ in this region since $x \in \mathcal{X}_0$.
- When $x > \lambda$, $\beta(x) = 1-(1-\alpha)\left(\frac{\lambda}{x}\right)^n$ for the same reasons as the previous region. The difference here is that $\beta(x) = P_D$ since $x \in \mathcal{X}_1$.

Putting this all together, we have

$$\beta(x) = \begin{cases} 1-(1-\alpha)(\frac{\lambda}{x})^n & \text{if } x \ge (1-\alpha)^{\frac{1}{n}}\lambda \\ 0 & \text{if } x < (1-\alpha)^{\frac{1}{n}}\lambda. \end{cases}$$

Figure 2c shows a plot of the power function for the case $\alpha = 0.1$, $\lambda = 1$, and $n = 1,\dots,5$. Note that the power function improves for increasing n and that the false positive probability constraint is satisfied for all $x \le \lambda$. Note that $(1-\alpha)^{\frac{1}{n}} \le 1$ for all

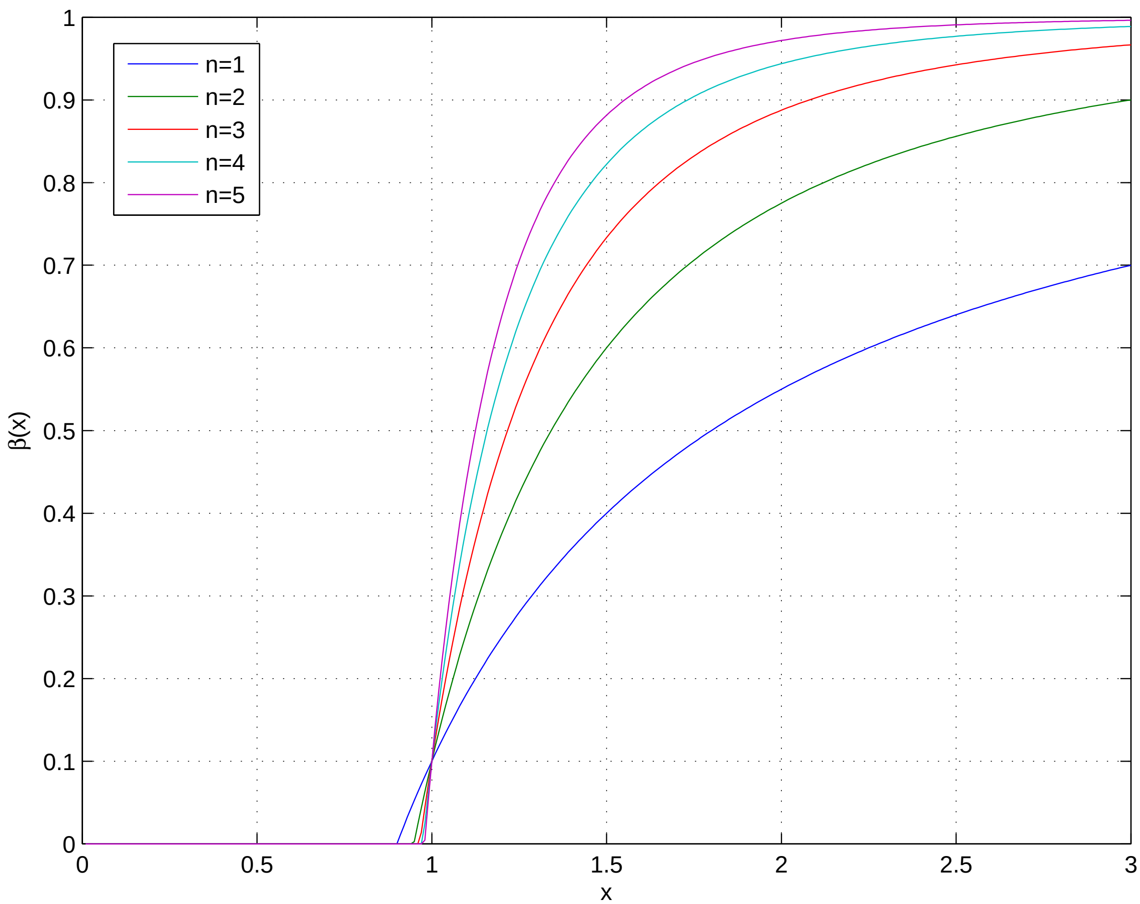

Figure 1: Problem 2c: Power function $\beta(x)$ versus x for the case $\alpha = 0.1$, $\lambda = 1$, and $n = 1,\dots,5$.

$\alpha \in [0,1]$. Hence, for all $x > \lambda$, i.e. all $x \in \mathcal{X}_1$, this decision rule has the same power as the decision rule considered in part (a). This confirms that the decision rule in part (a) is UMP. What is interesting, however, is that the UMP decision rule in part (c) has lower false positive probability than the decision rule in part (a). Both decision rules are UMP because they have the same power for all $x > \lambda$. But the decision rule in part (c) would be preferred because it has lower false positive probability for all $x \leq \lambda$.

3. 4 points. Consider a Laplacian random variable Y with unknown mean $x \in [0, \infty)$ and conditional density $p_x(y) = \frac{b}{2}e^{-b|y-x|}$. Given one observation of Y, we want to decide $\mathcal{H}_0 \leftrightarrow x = 0$ versus $\mathcal{H}_1 \leftrightarrow x > 0$ subject to an upper bound α on the false positive probability. Find a UMP decision rule and write an expression for its power function. You can check your results against the lecture notes.

   Solution: We know from the Lecture 5 notes that $p_x(y) = \frac{b}{2}e^{-b|y-x|}$ has monotone likelihood ratio for $x \in \mathcal{X}$. Thus by the Lehmann UMP theorem, and the fact that y has a continuous density, the decision rule

$$\rho(y) = \begin{cases} 1 & \text{if } y \geq \tau \\ 0 & \text{if } y < \tau \end{cases}$$

   is UMP where τ is selected so that $P_{\mathbf{fp},x=0}(\rho) = \alpha$.

   To determine τ, we compute the probability of a false positive when $x = 0$, i.e. we need to find τ such that

$$\int_\tau^\infty \frac{b}{2}e^{-b|y|}\,dy = \alpha.$$

   Note that $\alpha = 1/2$ implies $\tau = 0$. Hence, for $\alpha < 1/2$, we know that $\tau > 0$ and we can solve

$$\int_\tau^\infty \frac{b}{2}e^{-by}\,dy = \alpha.$$

   to determine that $\tau = -\ln(2\alpha)/b$. When $\alpha > 1/2$, we know that $\tau < 0$ and we can solve

$$\int_\tau^0 \frac{b}{2}e^{by}\,dy = \alpha - \frac{1}{2}$$

   to determine that $\tau = \ln(2(1-\alpha))/b$.

   The power function $\beta(x)$ is then computed by fixing $x > 0$ and computing the probability of correct detection

$$\beta(x) \quad = \quad \int_\tau^\infty \frac{b}{2}e^{-b|y-x|}\,dy.$$

   This has to be done on a case by case basis.

   - Case 1: $\alpha < 1/2$ such that $\tau = -\ln(2\alpha)/b > 0$. In this case, we have

$$\beta(x) \quad = \quad \begin{cases} \int_\tau^\infty \frac{b}{2}e^{-b(y-x)}\,dy & x \leq \tau \\ \int_\tau^x \frac{b}{2}e^{b(y-x)}\,dy + \int_x^\infty \frac{b}{2}e^{-b(y-x)}\,dy & x > \tau \end{cases}$$

     Note that the condition $x \leq \tau$ is equivalent to the condition $\alpha \leq \frac{1}{2}e^{-bx}$. Hence, when $0 \leq \alpha \leq \frac{1}{2}e^{-bx}$, we can write

$$\beta(x) \quad = \quad \int_\tau^\infty \frac{b}{2}e^{-b(y-x)}\,dy = \alpha e^{bx}$$

and when $\frac{1}{2}e^{-bx} < \alpha < \frac{1}{2}$, we can write

$$\beta(x) = \int_{\tau}^{x} \frac{b}{2} e^{b(y-x)}\, dy + \int_{x}^{\infty} \frac{b}{2} e^{-b(y-x)}\, dy = 1 - \frac{1}{4\alpha} e^{-bx}.$$

- Case 2: $\alpha \geq 1/2$ such that $\tau = \ln(2(1-\alpha))/b < 0$. In this case, we always have $x > \tau$. So we just need to compute

$$\beta(x) = \int_{\tau}^{x} \frac{b}{2} e^{b(y-x)}\, dy + \int_{x}^{\infty} \frac{b}{2} e^{-b(y-x)}\, dy = 1 - (1-\alpha)e^{-bx}.$$

The results are summarized in the following table

α	τ	$\beta(x)$
$0 \leq \alpha < \frac{1}{2}e^{-bx}$	$-\ln(2\alpha)/b$	αe^{bx}
$\frac{1}{2}e^{-bx} \leq \alpha < \frac{1}{2}$	$-\ln(2\alpha)/b$	$1 - \frac{1}{4\alpha}e^{-bx}$
$\frac{1}{2} \leq \alpha \leq 1$	$\ln(2(1-\alpha))/b$	$1-(1-\alpha)e^{-bx}$

Note that, when $x = 0$, $\beta(x) = \alpha$ in the first and third cases, which is a good check to make sure you haven't made a mistake somewhere. You don't need to worry about $x = 0$ in the second case because this case applies only when $x > -\ln(2\alpha)/b > 0$.

4. 6 points. Consider an observation of n i.i.d. Cauchy random variables

$$Y_i \sim \frac{1}{\pi(1+(y-\lambda)^2)}$$

and the one-sided test

$$\begin{aligned} \mathcal{H}_0 &: \lambda \leq 0 \\ \mathcal{H}_1 &: \lambda > 0 \end{aligned}$$

(a) Show that there is no UMP decision rule for this binary composite HT problem.

Solution: A UMP decision rule exists iff the critical region $\Gamma_1 = \{y \in \mathcal{Y} : \text{decide } \mathcal{H}_1\}$ is the same for all simple binary HT problems $\mathcal{H}_0 : \lambda = 0$ versus $\mathcal{H}_1 : \lambda = \lambda_1 > 0$. For the simple binary HT problem $\mathcal{H}_0 : \lambda = 0$ versus $\mathcal{H}_1 : \lambda = \lambda_1 > 0$ with n i.i.d. observations, we have the likelihood ratio

$$L(y) = \prod_{i=1}^{n} \frac{1+y_i^2}{1+(y_i-\lambda_1)^2}.$$

The procedure to find the critical region is to find the threshold v such that $\Gamma_1 = \{y \in \mathcal{Y} : L(y) > v\}$ and

$$\int_{y \in \Gamma_1} \prod_{i=1}^{n} \frac{1}{\pi(1+y_i^2)}\, dy = \alpha.$$

To see that the critical region depends on λ_1, hence a UMP detector can not exist, explicitly for the case when $n = 1$, we can write

$$\begin{aligned} & L(y) > v \\ \Leftrightarrow\ & \frac{1+y^2}{1+(y-\lambda_1)^2} > v \\ \Leftrightarrow\ & 1+y^2 > v + v(y-\lambda_1)^2 \\ \Leftrightarrow\ & 1+y^2 > v + vy^2 - 2vy\lambda_1 + v\lambda_1^2 \\ \Leftrightarrow\ & (1-v)y^2 + 2v\lambda_1 y + (1-v) - v\lambda_1^2 > 0 \end{aligned}$$

For sufficiently small α, v will be larger than one and the critical region will be an interval on the real line between the two distinct real roots of

$$y^2 - 2\frac{v\lambda_1}{v-1}y + \frac{v\lambda_1^2}{v-1} + 1 = 0.$$

You can use the quadratic formula to compute these roots analytically:

$$\begin{aligned} r_1 &= \frac{v\lambda_1}{v-1} - \sqrt{\frac{v\lambda_1^2}{(v-1)^2} - 1} \\ r_2 &= \frac{v\lambda_1}{v-1} + \sqrt{\frac{v\lambda_1^2}{(v-1)^2} - 1} \end{aligned}$$

hence the false positive probability can be computed as

$$\begin{aligned} \int_{r_1}^{r_2} \frac{1}{\pi(1+y^2)}\,dy &= \frac{1}{\pi}[\tan^{-1}(r_2) - \tan^{-1}(r_1)] \\ &= \frac{1}{\pi}\left[\tan^{-1}\left(\frac{v\lambda_1}{v-1} + \sqrt{\frac{v\lambda_1^2}{(v-1)^2} - 1}\right) - \tan^{-1}\left(\frac{v\lambda_1}{v-1} - \sqrt{\frac{v\lambda_1^2}{(v-1)^2} - 1}\right)\right] \end{aligned}$$

hence the problem is, given λ_1 and α, find v so that the false positive probability is equal to α. I don't think this can be done analytically, but it can be done numerically using `fsolve`, e.g.

```
[v,fval,exitflag] =
  fsolve(@(v) (1/pi)*(atan(v*lam1/(v-1)+sqrt(v*lam1^2/(v-1)^2-1))
             -atan(v*lam1/(v-1)-sqrt(v*lam1^2/(v-1)^2-1)))-alpha,2)
```

As an example, let $\alpha = 0.1$ and let $\lambda_1 = 1$. Matlab's `fsolve` function yields $v \approx 2.29$, which implies that $r_1 \approx 1.1619$ and $r_2 \approx 2.3885$, i.e. the critical region $\Gamma_1 = \{y : 1.1619 \leq y \leq 2.3885\}$. Now let's try $\alpha = 0.1$ and $\lambda_1 = 2$. In this case, Matlab's `fsolve` function yields $v \approx 3.2257$, which implies that $r_1 \approx 1.6318$ and $r_2 \approx 4.1653$, i.e. the critical region $\Gamma_1 = \{y : 1.6318 \leq y \leq 4.1653\}$. Note that the critical region has changed for two different values of λ_1, hence no UMP rule can exist.

(b) Find a LMP decision rule and its power function. Hint 1: You may need to numerically solve for the decision threshold. Hint 2: When n is large, you can use a Gaussian approximation on the test statistic to write the threshold as a Q function.

Solution: The LMP decision rule is based on $\frac{\partial L(y)}{\partial \lambda}$.

We have

$$L'_{\lambda=0}(y) = \sum_{i=1}^{n} \frac{2y_i}{1+y_i^2}$$

Using what you know about computing the mean of a function of a random variable, you can compute the mean of each term in the sum as

$$\mathrm{E}\left[\frac{2Y_i}{1+Y_i^2}\right] = \int_{-\infty}^{+\infty} \frac{1}{\pi}\frac{2y}{(1+y^2)^2}dy = 0$$

where this result follows immediately from the fact that the integrand is odd. The variance of each term in the sum can be computed as

$$\mathrm{E}\left[\left(\tfrac{2Y_i}{1+Y_i^2}\right)^2\right] = \textstyle\int_{-\infty}^{+\infty} \frac{1}{\pi} \frac{4y^2}{(1+y^2)^3} dy$$

$$= \tfrac{8}{\pi} \textstyle\int_0^{\pi/2} \cos^2 t \sin^2 t dt$$

$$= \tfrac{1}{\pi} \textstyle\int_0^{\pi/2} (1 - \cos 4t) dt$$

$$= \tfrac{1}{2}$$

Note that both the mean and variance of each term in the sum are finite. When n is large, the central limit theorem says that $L'_{\lambda=0}(y)$ will be approximately Gaussian. Hence, we can say that for large n

$$L'_{\lambda=0}(Y) = \sum_{i=1}^{n} \frac{2Y_i}{1+Y_i^2} = Z \sim \mathcal{N}(0, n/2).$$

and the LMP decision rule is

$$\rho(y) = \begin{cases} 1 & Z \geq \tau \\ 0 & Z < \tau \end{cases}$$

where randomization is not necessary because Z follows a continuous distribution. Thus

$$P_{\mathbf{fp}}(\rho) = \mathrm{Prob}[Z \geq \tau] = Q\left(\tau\sqrt{\frac{2}{n}}\right).$$

Setting this equal to α yields the threshold

$$\tau = \sqrt{\frac{n}{2}} Q^{-1}(\alpha)$$

which can be computed with Matlab's `erfcinv` function given n and α. The power function for $\lambda \geq 0$ is then

$$\beta(\lambda) \quad = \quad \mathrm{Prob}[Z \geq \tau \,|\, \text{state is } \lambda]$$

We need to know the statistics of Z when the state is $\lambda > 0$. We can compute the mean of each term in the sum as

$$\mathrm{E}\left[\frac{2Y_i}{1+Y_i^2}\right] = \int_{-\infty}^{+\infty} \frac{1}{\pi} \frac{2y}{(1+y^2)(1+(y-\lambda)^2)} dy = \frac{2\lambda}{\lambda^2+4}$$

and we can compute the variance of each term in the sum as

$$\begin{aligned} \mathrm{E}\left[\left(\frac{2Y_i}{1+Y_i^2}\right)^2\right] - \left(\frac{2\lambda}{\lambda^2+4}\right)^2 &= \int_{-\infty}^{+\infty} \frac{1}{\pi} \frac{4y^2}{(1+y^2)(1+(y-\lambda)^2)^2} dy \\ &= \frac{2\lambda^4 + 10\lambda^2 + 8}{(\lambda^2+4)^2} \\ &= \frac{2(\lambda^2+1)}{\lambda^2+4} \end{aligned}$$

where I used Matlab's symbolic toolbox to compute these integrals. Note that these agree with the prior results when $\lambda = 0$. Using the same reasoning as before, we can say that $Z \sim \mathcal{N}\left(\mu, \sigma^2\right)$ with

$$\begin{aligned}
\mu &= \frac{2n\lambda}{\lambda^2 + 4} \\
\sigma^2 &= \frac{2n(\lambda^2 + 1)}{\lambda^2 + 4}
\end{aligned}$$

hence the power function can be written as

$$\beta(\lambda) = \mathrm{Prob}[Z \geq \tau \,|\, \text{state is } \lambda] \approx Q\left(\frac{\tau - \mu}{\sigma}\right)$$

where τ, μ, and λ are all given above.