Solutions to the Exercises of Section 2.11.

2.11.1. **Proof.** Let ϵ be an arbitrary positive number. Since $r(\tau_n,\delta_n)\to C$, we can find an integer n such that $r(\tau_n,\delta_n)\ge C-\epsilon$. Then, as in the proof of Theorem 1,

$$\overline{V}\le \sup_\theta R(\theta,\delta_0)\le C\le r(\tau_n,\delta_n)+\epsilon\le \inf_\delta r(\tau_n,\delta)+\epsilon\le \underline{V}+\epsilon.$$

Since this holds for all $\epsilon>0$, we have

$$\overline{V}\le \sup_\theta R(\theta,\delta_0)\le C\le \underline{V}.$$

This shows that the value exists and is equal to C and that δ_0 is minimax. ∎

2.11.2. **Proof.** Let τ_0 be least favorable and let θ_0 be in the support of τ_0. (This means for every $\epsilon>0$ that $\tau_0(\theta_0-\epsilon,\theta_0+\epsilon)>0$.) Suppose that $R(\theta_0,\delta_0)<V$. Since $R(\theta,\delta_0)$ is continuous, there exists an $\epsilon>0$ such that $R(\theta,\delta_0)<V$ for all θ in $(\theta_0-\epsilon,\theta_0+\epsilon)$, a set of positive measure. But since δ_0 is minimax, we must have $R(\theta,\delta_0)\le V$ for all θ. Hence,

$$V=r(\tau_0,\delta_0)=\int R(\theta,\delta_0)\,d\tau_0(\theta)<V.$$

This contradiction completes the proof. ∎

2.11.3. **Proof:** Let δ_0 be an equalizer rule with $R(\theta,\delta_0)\equiv c$. If δ_0 were not minimax, then there would exist a rule δ' such that $\sup_\theta R(\theta,\delta')$, call it v, would be strictly less than $\sup_\theta R(\theta,\delta_0)=c$. For any ϵ such that $0<\epsilon<c-v$ and any prior, π, we have

$$r(\pi,\delta_0)=c>v+\epsilon\ge r(\pi,\delta')+\epsilon.$$

Thus, δ_0 cannot be ϵ-Bayes with respect to any prior distribution, and so δ_0 cannot be an extended Bayes rule. ∎

2.11.4. If d is an equalizer rule, then $R(\theta,d)\equiv c$ for some constant c. Hence, $r(\tau,d)=c$ for all $\tau\in\Theta^*$, and hence $\sup_\tau r(\tau,d)=c$. If d were not minimax, then there would exist a rule d_0 with smaller maximum risk: $\sup_\tau r(\tau,d_0)=c_1<c$. This implies $R(\theta,d_0)\le c_1<c=R(\theta,d)$ for all θ, showing that d is not admissible.

2.11.5. Take $\Theta=\{1,2,\ldots\}$ and $\mathcal{A}=\{0,1,2,\ldots\}$ and $L(\theta,a)=1$ if $\theta<a$, $L(\theta,a)=-1$ if $0<a<\theta$, and $L(\theta,a)=0$ if $a=\theta$ or if $a=0$. Then $d\equiv 0$ is an equalizer rule ($L(\theta,0)\equiv 0$), and d is admissible since nature can do better than zero against any $\delta\ne d$ by taking θ large enough. Hence, d is minimax from 2.11.4. But the value does not exist since the lower value is still -1 as in (2.24) on pg. 83, and the upper value is now zero. Furthermore, d cannot be Bayes with respect to any τ, since $r(\tau,d)\equiv 0$ and the statistician can do better than zero against τ by choosing a sufficiently large.

2.11.6. In the problem with $\Theta=\{\theta_1,\theta_2\}$, $\mathcal{A}=[0,\pi/2]$ and loss $L(\theta_1,a)=-\cos(a)$, $L(\theta_2,a)=-\sin(a)$, the loss is convex in a for each θ so that attention may be restricted to the nonrandomized rules. The observations are $X=0$ or 1, with $P(X=1|\theta_1)=1/3$, and $P(X=1|\theta_2)=2/3$. The nonrandomized rules are of the form (a_0,a_1) where a_j is the action we take if $X=j$ is observed, $j=0$ or 1. For $d=(a_0,a_1)$, we have

$$R(\theta_1,d)=-(1/3)\cos(a_1)-(2/3)\cos(a_0)$$
$$R(\theta_2,d)=-(2/3)\sin(a_1)-(1/3)\sin(a_0).$$

If τ_0 gives probability $1/2$ to each state of nature, then

$$r(\tau_0,d)=-(1/6)(\cos(a_1)+2\sin(a_1)+2\cos(a_0)+\sin(a_0)).$$

The values of a_0 and a_1 that minimize this may be found by taking derivatives and equating to zero. This gives $\tan(a_1) = 2$ and $\tan(a_0) = 1/2$, which gives the Bayes rule with respect to τ_0, call it d_0. From this, we find $\sin(a_1) = 2/\sqrt{5} = \cos(a_0)$ and $\cos(a_1) = 1/\sqrt{5} = \sin(a_0)$, so that the risk function of d_0 is

$$\begin{aligned} R(\theta_1, d_0) &= -(1/3)(1/\sqrt{5} + 4/\sqrt{5}) = -\sqrt{5}/3. \\ R(\theta_2, d_0) &= -(1/3)(4/\sqrt{5} + 1/\sqrt{5}) = -\sqrt{5}/3. \end{aligned}$$

Thus, the Bayes risk of d_0 is also $-\sqrt{5}/3$, so from Theorem 2.11.1 we may conclude that d_0 is minimax and τ_0 is least favorable.

2.11.7. (a) Since $\mathcal{A} = [0,1]$ and X takes on two values, say $X = 1$ if heads and $X = 0$ if tails, D is the unit square, where $d = (a_0, a_1)$ represents the decision to choose a_j if $X = j$.

(b) If $d = (x, y) \in D$, then

$$\begin{aligned} R(1/3, d) &= (2/3)x^2 + (1/3)y^2 \\ R(2/3, d) &= (1/3)(1-x) + (2/3)(1-y). \end{aligned}$$

(c) A prior distribution may be represented by p, the probability that $\theta = 1/3$. The Bayes risk of $d \in D$ is

$$r(p, d) = p[2x^2 + y^2]/3 + (1-p)[(1-x) + 2(1-y)]/3.$$

Setting derivatives with respect to x and y to zero gives $4px - (1-p) = 0$ and $2py - 2(1-p) = 0$, from which we find the Bayes rule to be

$$\begin{array}{llllll} x = (1-p)/4p & \text{if } p > 1/5 & \text{and} & x = 1 & \text{if} & p > 1/5. \\ y = (1-p)/p & \text{if } p > 1/2 & \text{and} & y = 1 & \text{if} & p \le 1/2. \end{array}$$

The set of such (x, y) as p ranges from 0 to 1 consists of two lines, the line $y = 4x$ from (0,0) to (1/4,1) (for $p > 1/2$), and the line $y = 1$ from (1/4,1) to (1,1) (for $p \le 1/2$).

(d) We search for an equalizer rule on the line $y = 4x$. Set $R(1/3, (x, 4x)) = R(2/3, (x, 4x))$ and solve for x: $(2/3)x^2 + (1/3)(4x)^2 = (1/3)(1-x) + (2/3)(1-4x)$ implies that $6x^2 + 3x - 1 = 0$. Since this gives $x = (\sqrt{33} - 3)/12 = .2287\cdots$ and $y = (\sqrt{33} - 3)/3 = .9148\cdots$ and both are in $\mathcal{A}$, this is the minimax rule. The minimax risk is $1 - 3x = .3139\cdots$.

2.11.8. The risk function of the rule $d(x) = x/n$ is

$$\begin{aligned} R(\theta, d) &= E\{(X/n - \theta)^2/(\theta(1-\theta))|\theta\} = \mathrm{Var}(X/n|\theta)/(\theta(1-\theta)) \\ &= \{\theta(1-\theta)/n\}/(\theta(1-\theta)) = 1/n, \end{aligned}$$

a constant, so that d is an equalizer rule. From Exercise 1.8.9, d is a Bayes rule for the uniform prior distribution on θ. Therefore, from Theorem 3, the rule d is minimax.

2.11.9. (a)

$$\begin{aligned} R(\theta, d) &= E_\theta\left(\frac{(\theta - d(X))^2}{(1-\theta)}\right) \\ &= \sum_{x=0}^{\infty} (\theta^2 - 2\theta d(x) + d(x)^2)\theta^x \\ &= d(0)^2 + \theta(d(1)^2 - 2d(0)) + \sum_{x=2}^{\infty} (d(x)^2 - 2d(x-1) + 1)\theta^x \end{aligned}$$

(b) d is an equalizer rule if and only if

$$\begin{aligned} & d(1)^2 = 2d(0) \qquad \text{and} \\ & d(x)^2 = 2d(x-1) - 1 \quad \text{for} \quad x = 2, 3, \ldots \end{aligned}$$

We are to show that the only solution in $\mathcal{A}$ to these equations is given by $d(0) = 1/2$ and $d(x) = 1$ for $x \geq 1$. For $x \geq 2$, the lower equations imply that $2d(x-1) = d(x)^2 + 1 = (d(x)-1)^2 + 2d(x) \geq 2d(x)$, so that the $d(x)$ are nonincreasing from $x = 1$ on. Hence, the d(x) converge, say, to some number c, and applying $\lim_{x\to\infty}$ to the lower equations gives $c^2 = 2c - 1$, from which we conclude that c must be equal to one. The only nonincreasing sequence in [0,1] that converges to one is the sequence identically one. Hence, $d(1) = d(2) = \cdots = 1$, and from the top equation, $d(0) = 1/2$ is the unique equalizer rule.

(c) From part (a),

$$\begin{aligned} r(\tau, d) &= E\,R(\theta, d) \\ &= d(0)^2 + \mu_1(d(1)^2 - 2d(0)) + \sum_{x=2}^{\infty}(d(x)^2 - 2d(x-1) + 1)\mu_x \end{aligned}$$

We may find the Bayes rule by taking the derivative with respect to each $d(x)$ separately, and setting equal to zero, to find

$$d_\tau(x) = \mu_{x+1}/\mu_x \quad \text{for} \quad x = 0, 1, 2, \ldots$$

(d) For d to be a Bayes rule with respect to τ, we must have $d(x) = \mu_{x+1}/\mu_x$ for $x = 0, 1, 2, \ldots$, which reduces to $\mu_1 = \mu_2 = \mu_3 = \cdots = 1/2$. There is a distribution, τ, with these moments, namely, the two point distribution, $\tau(0) = 1/2$, and $\tau(1) = 1/2$. Unfortunately, 1 is not in the parameter space, so d is not a Bayes rule. To show that it is extended Bayes, we look at the distributions, τ_ϵ, that give weight 1/2 to 0 and weight 1/2 to $1-\epsilon$. For this distribution, $\mu_x = (1-\epsilon)^x/2$ and the Bayes rule with respect to τ_ϵ is $d_\epsilon(0) = (1-\epsilon)/2$, and $d_\epsilon(x) = 1-\epsilon$ for $x = 1, 2, \ldots$. To compute the minimum Bayes risk, note since d_ϵ takes on only two values,

$$R(\theta, d_\epsilon) = (1-\theta)\frac{(\theta - (1-\epsilon)/2)^2}{1-\theta} + \theta\frac{(\theta - (1-\epsilon))^2}{1-\theta}$$

and since θ takes on only two values,

$$r(\tau_\epsilon, d_\epsilon) = \frac{1}{2}(\frac{1-\epsilon}{2})^2 + \frac{1}{2}(\epsilon - \frac{1-\epsilon}{2})^2 = \frac{(1-\epsilon)^2}{4}.$$

Since the constant risk of the rule d is 1/4, and since the minimum Bayes risk of τ_ϵ converges to 1/4 as $\epsilon \to 0$, d is extended Bayes and hence minimax.

2.11.10. (a) The risk function of the rule d_0 is

$$\begin{aligned} R(\theta, d_0) &= E_\theta(\mu_1 - \frac{X+\sqrt{n}}{n+\sqrt{n}})^2 = \mathrm{Var}(\frac{X+\sqrt{n}}{n+\sqrt{n}}) + (\mu_1 - \frac{n\mu_1+\sqrt{n}}{n+\sqrt{n}})^2 \\ &= \frac{n\sigma^2}{(n+\sqrt{n})^2} + \frac{(\mu_1\sqrt{n} - \sqrt{n}/2)^2}{(n+\sqrt{n})^2} = \frac{n(\mu_2 - \mu_1 + 1/4)}{(n+\sqrt{n})^2} \end{aligned}$$

This risk is maximized by that distribution, θ on $[0,1]$, that maximizes $\mu_2 - \mu_1 = -E(Z(1-Z))$, where Z has distribution θ. This is never positive and is equal to zero if and only if θ gives all its mass to the points $Z = 0$ and $Z = 1$. The maximum risk is then $\max_\theta R(\theta, d_0) = (1/4)/(\sqrt{n}+1)^2$.

(b) From pages 93-94, d_0 is a Bayes rule with respect to the distribution that chooses p at random from a Beta distribution, $\mathcal{B}e(\sqrt{n}/2, \sqrt{n}/2)$, and gives mass p to 1 and mass $1-p$ to 0. The minimum Bayes risk with respect to this prior is $(1/4)/(\sqrt{n}+1)^2$, the same as the maximum risk of the rule d_0. Hence d_0 is minimax from Theorem 1.

2.11.11. (a) $R(\theta, d_0) = E_\theta\{(\theta - X)^2/\theta\} = \theta/\theta = 1$.

(b) The generalized Bayes risk is

$$\sum_{x=0}^{\infty}\int_0^\infty (\theta - d(x))^2 e^{-\theta}\theta^{x-1}/x!\,d\theta.$$

For $x = 0$, the integral is $+\infty$ unless $d(0) = 0$. For $x > 0$, the integral is minimized if $d(x)$ is chosen to be the mean of the gamma distribution, $\mathcal{G}(x, 1)$, namely, $d(x) = x$. Thus, d_0 is a generalized Bayes rule.

(c) For the prior $\tau_{\alpha,\beta} = \mathcal{G}(\alpha, \beta)$, the posterior distribution is proportional to $e^{-\theta}\theta^x e^{-\theta/\beta}\theta^{\alpha-1}$, so that the posterior distribution is $\mathcal{G}(\alpha + x, \beta/(\beta + 1))$. The Bayes risk is proportional to

$$\sum_{x=0}^{\infty} \int_0^{\infty} (\theta - d(x))^2 e^{-\theta-\theta/\beta}\theta^{\alpha+x-2}\, d\theta/x!.$$

If $\alpha + x - 1 \le 0$, the integral is infinite unless $d(0) = 0$. For $\alpha + x - 1 > 0$, the integral is minimized by the mean of $\mathcal{G}(\alpha + x - 1, \beta/(\beta + 1))$, namely, $(\alpha + x - 1)\beta/(\beta + 1)$. This determines the Bayes rule to be

$$d_{\alpha,\beta}(x) = \max\{0, (\alpha + x - 1)\beta/(\beta + 1)\}.$$

(d) Take $\alpha = 1$ and find the risk function of the rule $d_{1,\beta}$

$$\begin{aligned} R(\theta, d_{1,\beta}) &= E_\theta(\theta - \beta X/(\beta + 1))^2/\theta \\ &= \mathrm{Var}(\beta X/(\beta + 1)|\theta)/\theta + (\theta - \beta\theta/(\beta + 1))^2/\theta \\ &= (\beta/(\beta + 1))^2 + \theta/(\beta + 1)^2. \end{aligned}$$

The minimum Bayes risk is thus

$$\begin{aligned} r(\tau_{1,\beta}, d_{1,\beta}) &= ER(\theta, d_{1,\beta}) \\ &= (\beta/(\beta + 1))^2 + \beta/(\beta + 1)^2 = \beta/(\beta + 1). \end{aligned}$$

As $\beta \to \infty$, this risk tends to 1. Since the rule of part (a) has constant risk 1, its Bayes risk is also 1, showing that it is ϵ-Bayes for every $\epsilon > 0$. This implies that it is minimax.

2.11.12. (a) X has the negative binomial distribution, $\mathcal{NB}(r, p) = \mathcal{NB}(r, \theta/(\theta+1))$, where $\theta = p/(1-p)$ represents the odds. Since $E(X|\theta) = rp/(1 - p) = r\theta$ and $\mathrm{Var}(X|\theta) = rp/(1 - p)^2 = r\theta(\theta + 1)$, we have

$$R(\theta, d_0) = \mathrm{Var}(X/r|\theta)/(\theta(\theta + 1)) = r\theta(\theta + 1)/(r^2\theta(\theta + 1)) = 1/r.$$

(b) The generalized Bayes rule minimizes for each x

$$\int_0^{\infty} \frac{(\theta - d(x))^2}{\theta(\theta + 1)}\, \theta^x(\theta + 1)^{-(r+x)}\, d\theta. \tag{1}$$

If $x = 0$, the integral is infinite unless $d(0) = 0$. For $x > 0$, the minimum occurs at

$$\begin{aligned} d(x) &= \int \theta^x(\theta + 1)^{-(r+x+1)}\, d\theta \Big/ \int \theta^{x-1}(\theta + 1)^{-(r+x+1)}\, d\theta \\ &= \int p^x(1 - p)^{r-1}\, dp \Big/ \int p^{x-1}(1 - p)^r\, dp \\ &= (\Gamma(x + 1)\Gamma(r)/\Gamma(x + r + 1))/(\Gamma(x)\Gamma(r + 1)/\Gamma(x + r + 1)) = x/r. \end{aligned}$$

(A minor point: If $r = 1$, the integral (1) is infinite no matter what $d(x)$ is chosen to be, so technically $d(x)$ could be anything.)

(c) When θ has the indicated distribution ($p \in \mathcal{B}\mathrm{e}(\alpha, \beta)$), the minimum of the Bayes risk occurs at 0 if $\alpha + x - 1 \le 0$, and otherwise at

$$d_{\alpha,\beta}(x) = \frac{\int \theta^{\alpha+x-1}(\theta + 1)^{-(\alpha+\beta+r+x+1}\, d\theta}{\int \theta^{\alpha+x-2}(\theta + 1)^{-(\alpha+\beta+r+x+1}\, d\theta} = \frac{\alpha + x - 1}{\beta + r + 1}.$$

(d) The risk function of $d(x) = x/(r+1)$ is

$$\begin{aligned} R(\theta, d) &= [\mathrm{Var}(X/(r+1)|\theta) + (\theta - r\theta/(r+1))^2]/(\theta(\theta+1)) \\ &= r/(r+1)^2 + (\theta/(\theta+1))/(r+1)^2. \end{aligned}$$

Since $\theta/(\theta+1) < 1$ for all θ, we have $R(\theta, d) < 1/(r+1) < 1/r$, the risk of d_0, so d_0 is not minimax in spite of the fact that it is an equalizer and generalized Bayes. (This shows that d_0 cannot be extended Bayes.) To show that d is minimax, we note that d is a limit of Bayes rules in part (c) for $\alpha = 1$ as $\beta \to 0$. We show that the minimum Bayes risk of $d_{1,\beta}$ tends to $1/(r+1)$ as $\beta \to 0$; then, Theorem 2 implies that d is minimax.

$$\begin{aligned} r(\tau_{1,\beta}, d_{1,\beta}) &= E\{(\theta - X/(\beta+r+1))^2/(\theta(\theta+1))\} \\ &= E\{[\mathrm{Var}(X/(\beta+r+1)|\theta) + (\theta - r\theta/(\beta+r+1))^2]/(\theta(\theta+1))\} \\ &= E\{r/(\beta+r+1)^2 + ((\beta+1)/(\beta+r+1))^2(\theta/(\theta+1))\} \\ &= r/(\beta+r+1)^2 + ((\beta+1)/(\beta+r+1))^2(1/(1+\beta)) \\ &\to r/(r+1)^2 + 1/(r+1)^2 \\ &= 1/(r+1). \end{aligned}$$

2.11.13. The rule d_0 of formula(2.32) is admissible for the problem considered there because it is unique Bayes with respect to the $\mathcal{B}e(\sqrt{n}/2, \sqrt{n}/2)$ prior.

2.11.14. (a) The density functions for θ and X are

$$g(\theta) = \binom{M}{\theta} \frac{\mathrm{B}(\alpha+\theta, \beta+M-\theta)}{\mathrm{B}(\alpha,\beta)} \qquad\qquad f(x|\theta) = \frac{\binom{n}{x}\binom{M-n}{\theta-x}}{\binom{M}{\theta}},$$

where $\mathrm{B}(\alpha,\beta)$ is the beta function, $\mathrm{B}(\alpha,\beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$. Since $\sum g(\theta) = 1$, we have the identity

$$\sum_\theta \binom{M}{\theta} \mathrm{B}(\alpha+\theta, \beta+M-\theta) = \mathrm{B}(\alpha,\beta)$$

from which we may compute the marginal distribution of X:

$$\begin{aligned} f(x) &= \sum_\theta \binom{n}{x}\binom{M-n}{\theta-x} \frac{\mathrm{B}(\theta+\alpha, M+\beta-\theta)}{\mathrm{B}(\alpha,\beta)} \qquad\qquad (\text{let } \tau = \theta - x) \\ &= \binom{n}{x} \sum_\tau \binom{M-n}{\tau} \frac{\mathrm{B}(\tau+x-\alpha, (M-n)+(n-x+\beta))}{\mathrm{B}(\alpha,\beta)} \\ &= \binom{n}{x} \frac{\mathrm{B}(\alpha+x, \beta+n-x)}{\mathrm{B}(\alpha,\beta)} \end{aligned}$$

which is beta-binomial, $\mathcal{BB}(\alpha,\beta,n)$.

(b) The conditional distribution of $\tau = \theta - x$ given $X = x$ has density $f(x|\tau+x)g(\tau+x)/f(x)$ which is easily computed to be the beta-binomial, $\mathcal{BB}(\alpha+x, \beta+n-x, M-n)$.

(c) Since the mean of $\mathcal{BB}(\alpha,\beta,n)$ is $n\alpha/(\alpha+\beta)$, the Bayes estimate of θ, being x plus the mean of $\mathcal{BB}(\alpha+x, \beta+n-x, M-n)$, is

$$\begin{aligned} d_{\alpha,\beta}(x) &= x + (M-n)(\alpha+x)/(\alpha+\beta+n) \\ &= [(M+\alpha+\beta)x + \alpha(M-n)]/(\alpha+\beta+n) \end{aligned}$$

which is linear in x.

(d) The risk function of the linear estimate $d(x) = ax + b$ is

$$\begin{aligned} R(\theta, d) &= E_\theta(\theta - aX - b)^2 \\ &= \mathrm{Var}(aX + b|\theta) + (\theta - an\theta/M - b)^2 \\ &= a^2 n\theta(M - \theta)(M - n)/(M^2(M - 1)) + (\theta(1 - an/M) - b)^2 \end{aligned}$$

(e) Assume $0 < n < M$ to avoid trivial cases. The risk in (d) is quadratic in θ and can be made constant by choosing a and b so that the linear and quadratic terms disappear, giving constant risk b^2. The linear and quadratic terms are proportional to

$$\theta^2[-a^2 n(M - n) + (M - an)^2(M - 1)] \qquad \text{and} \qquad \theta[a^2 n(M - n) - 2b(M - an)(M - 1)]$$

Setting the coefficients of the linear and quadratic terms to zero gives $a = M/2$, $b = M/4$ when $n = 1$, and for $n > 1$:

$$\begin{aligned} &a = (1 \pm \delta)(M - 1)/(n - 1) \qquad \text{and} \\ &b = n(M - n)a^2/(2(M - an)(M - 1)) = (M - an)/2 \end{aligned}$$

(or the values given in the text), where $\delta = ((M - n)/n(M - 1))^{1/2}$.

(f) To check that $d(x)$ is a Bayes rule, we equate to (c) and solve for α and β to see that both are positive. This occurs if the minus sign is used in the formula for a, and we find

$$\alpha = \beta = b/(a - 1) = M\delta n/(2(M - (1 + \delta)n).$$

This is positive if and only if $n < M - 1$, and so the resulting rule is minimax in this case. For example, if $M = 9$ and $n = 3$, we find $\delta = 1/2$, $a = 2$, and $b = 3/2$, so that $d(x) = 2x + 3/2$ is minimax with constant risk 9/4.

If $n = M - 1$, the above analysis does not work, but one can try to show that the corresponding equalizer $d(x) = x + 1/2$ is minimax by showing it is extended Bayes (ϵ-Bayes with respect to $\mathcal{BB}(\alpha, \beta, M)$ for $\alpha = \beta$ sufficiently large) or Bayes with respect to the binomial distribution, $\mathcal{B}(M, 1/2)$.

2.11.15. The rule $d(x) \equiv 1/2$ has constant risk function, $R(\theta, d) \equiv 1$. Moreover, the risk at $\theta = 1/2$ of any rule, δ, is 1, $R(1/2, \delta) = 1$, so $\inf_\delta \sup_\theta R(\theta, \delta) = 1$. Hence, $\sup_\theta R(\theta, d) = \inf_\delta \sup_\theta R(\theta, \delta)$, and d is minimax.

2.11.16. The risk function for the rule $d = \mathbf{x} \in D = \{\mathbf{x} = (x_0, x_1, \cdots, x_n) : 0 \le x_j \le 1 \text{ for all } j\}$ is

$$R(\theta, d) = \sum_{j=0}^{n} \binom{n}{j} \theta^j (1 - \theta)^{n-j} L(\theta, x_j).$$

Since $L(\theta, 0) \equiv 1$, the rule $d_0 = \mathbf{0}$ has constant risk, $R(\theta, \mathbf{0}) = 1$. We are to show that d_0 is the unique minimax rule. We will show that for any rule $\delta \in D^*$ other than d_0 there exists a value of θ such that $R(\theta, \delta) > 1$.

$$\begin{aligned} R(\theta, \delta) &= \int R(\theta, \mathbf{x})\, d\delta(\mathbf{x}) \\ &= \sum_{j=0}^{n} \binom{n}{j} \theta^j (1 - \theta)^{n-j} L(\theta, \delta_j) \end{aligned}$$

where $\delta_j \in \mathcal{A}^*$ is the marginal distribution of x_j. First we note the following lemma.

Lemma. *For* $\delta \in \mathcal{A}^*$, $L(\theta, \delta) \to 1 + \delta(0, 1]$ *as* $\theta \to 0$.

Proof.

$$\begin{aligned} L(\theta, \delta) &= \int L(\theta, a)\, d\delta(a) = \int_{0^+}^{1} L(\theta, a)\, d\delta(a) + \delta(0) \\ &\to \int_{0^+}^{1} 2\, d\delta(a) + \delta(0) \\ &= 2\delta(0, 1] + 1 - \delta(0, 1] = 1 + \delta(0, 1]. \ \blacksquare \end{aligned}$$

We now proceed to show that for any rule δ not degenerate at $\mathbf{0}$, $R(\theta, \delta) > 1$ for θ sufficiently close to zero. If, δ_0 is not degenerate at zero, then $R(\theta, \delta) \to 1 + \delta_0(0, 1] > 1$ as $\theta \to 0$. If δ_0 is degenerate at zero, then $R(\theta, \delta) \to 1$, unfortunately; so instead we work with

$$R(\theta, \delta) - 1 = \sum_{j=0}^{n} \binom{n}{j} \theta^j (1-\theta)^{n-j} (L(\theta, \delta_j) - 1)$$

If δ_0 is degenerate at zero but δ_1 is not, then

$$\begin{aligned}(R(\theta, \delta) - 1)/\theta &= n(1-\theta)^{n-1}(L(\theta, \delta_1) - 1) + \theta \cdot (\text{other terms}) \\ &\to \delta_1(0, 1] > 0 \qquad \text{as} \qquad \theta \to 0.\end{aligned}$$

Hence, $R(\theta, \delta) > 1$ for θ sufficiently close to zero. Similarly, if δ is not degenerate at $\mathbf{0}$, there is a smallest j such that δ_j is not degenerate at zero, and then

$$(R(\theta, \delta) - 1)/\theta^j \to \delta_j(0, 1] > 0,$$

showing that $R(\theta, \delta) > 1$ for θ sufficiently close to zero.