Decision Rules

A *nonrandomized decision rule* is a function d that assigns a specific action to each experimental outcome, i.e., $d : \mathcal{X} \to \mathcal{A}$. Let $\mathcal{D}$ denote the class of all nonrandomized decision rules. The risk set of $\mathcal{D}$ is the set of risk functions that correspond to elements of $\mathcal{D}$.

There are two standard ways to randomize decisions.

- A *randomized decision rule* is a probability distribution on $\mathcal{D}$. Let $\mathcal{D}^*$ denote the class of all randomized decision rules and note that nonrandomized decision rules are special cases of randomized decision rules, i.e., $\mathcal{D} \subset \mathcal{D}^*$. Accordingly, we often refer to $\mathcal{D}^*$ as the class of decision rules.[1]

  If $\delta \in \mathcal{D}^*$ is the decision rule and we observe x, then we proceed as follows. First draw $d \sim \delta$, then take action $d(x)$. The risk function that corresponds to δ is

  $$R(\theta, \delta) = E_\delta R(\theta),$$

  where Z is a random element that assumes values in D according to probability distribution δ.

  The risk set of $\mathcal{D}^*$, i.e., the set of risk functions that correspond to elements of $\mathcal{D}^*$, is the convex hull of the risk set of $\mathcal{D}$.

- A *behavioral decision rule* is a function δ that assigns a probability distribution on $\mathcal{A}$ to each experimental outcome, i.e., $\delta : \mathcal{X} \to \mathcal{A}^*$. Using a behavioral decision rule, if we observe x, then we take action $a \sim \delta(x)$. The loss function of a behavioral decision rule is

  $$L(\theta, \delta(x)) = E_{\delta(x)} L(\theta, a)$$

  and the corresponding risk function is

  $$R(\theta, \delta) = E_\theta L(\theta, \delta(X)).$$

It turns out that the classes of randomized and behaioral decison rules are equivalent. It is fairly obvious that the risk set of the behavioral decision rules contains the risk set of $\mathcal{D}^*$. To see this, note that randomization is performed *after* we observe x. Suppose that $\delta \in \mathcal{D}^*$. Then we take action $a \in \mathcal{A}$ according to the probability that drawing $d \sim \delta$ produces a nonrandomized decision rule for which $d(x) = a$. Thus, $\delta \in \mathcal{D}^*$ induces a probability distribution on $\mathcal{A}$, which is precisely what we mean by a behavioral decision rule.

The converse statement, that the risk set of $\mathcal{D}^*$ contains the risk set of the behavioral decision rules, is also true but considerably more difficult to prove—well beyond the scope of this course. Depending on what one is trying to do, one may wish to work with either randomized or behavioral decision rules.

[1] Here I am mimicking the notation of T. S. Ferguson's *Mathematical Statistics: A Decision Theoretic Approach*, 1967, a standard reference for statistical decision theory. Note that Young & Smith use $\mathcal{D}$ to denote the class of all randomized decision rules.