Chapter 7

Model Uncertainty and Misspecification

We have so far assumed that agents' model of the world is *correctly specified*: Their prior belief over Θ assigns positive probability to the true parameter θ and they update to information correctly, i.e. with knowledge of the true signal generating distribution $(P_\theta)_{\theta \in \Theta}$. Some reasons to question this model of learning include:

- We see substantial and persistent disagreement between individuals, but Sections 6.3 and 6.4 imply that agents will eventually hold similar beliefs.
- It is unclear how agents came to know $(P_\theta)_{\theta \in \Theta}$.
- The assumption that agents perceive only one signal-generating distribution $(P_\theta)_{\theta \in \Theta}$ as possible means that agents never abandon their model, even as evidence accumulates against it. As we discuss in Section 7.1.1, this dogmatism has some strange implications.

This section relaxes the standard learning model by allowing for *model uncertainty* (Section 7.1) and *model misspecification* (Section 7.2). In the former class of models, agents hold non-degenerate beliefs over the signal generating distribution. In the second, agents assign probability zero to the true parameter.

7.1 Model Uncertainty

7.1.1 Motivation

Recall the binary setting from Section 6.2: There is an unknown parameter $\theta \in \{A, B\}$, and each period $t \in \mathbb{Z}_+$ a signal is generated iid according to

	a	b
A	q	$1-q$
B	$1-q$	q

where $q > 1/2$. Agents may hold different (non-degenerate) prior beliefs $\pi_i \in \Delta(\Theta)$ about the parameter, but the value of q is common knowledge.

In Section 6.2, we observed that these agents almost surely learn the true parameter as the sample size grows large, and moreover their disagreement about the parameter vanishes. This is because (1) agents assign probability 1 to the event in which the limiting fraction of a-realizations is either q or $1-q$, and (2) the parameter is identified, so for either of these limiting frequencies agents (eventually) assign probability 1 to the correct parameter value.

What happens along sequences in which the limiting frequency is neither $(q, 1-q)$ nor $(1-q, q)$? Although agents assign probability zero to this event, sampling variation can explain any empirical frequency of a and b realizations (however surprising) in finite sequences. Thus Bayes' rule yields well-defined posterior beliefs.

For example, suppose $q \in (1/2, 1)$ and let $\mathbf{x}$ be the (infinite) sequence of a-realizations. For any t, the unconditional probability of the event that all t realizations are a is

$$\pi^i_A \cdot q^t + (1-\pi^i_A) \cdot (1-q)^t$$

where π^i_A denotes the prior probability of A. This expression converges to zero as t grows large but is strictly positive for every t. The agent's limiting belief along $\mathbf{x}$ can thus be computed to be

$$\lim_{t\to\infty} P^i(\theta = A \mid \mathbf{x}_t) = \lim_{t\to\infty} \frac{1}{1 + \frac{1-\pi^i_A}{\pi^i_A}\left(\frac{1-q}{q}\right)^t} = 1$$

So the agent is increasingly convinced that the state is A, even as the observed sequence grows increasingly unlikely under the agent's model. Even more striking, as signals accumulate in the frequency $(1, 0)$, the agent becomes increasingly confident that future signals will appear in the frequency $(q, 1-q)$! These conclusions are a consequence of the agent's dogmatic view of the signal generating distribution—he is unwilling to abandon this model even as mounting evidence points to its error.

7.1.2 Expanded Framework

We can introduce *model uncertainty* into this learning model by expanding the state space to $\Omega = \Theta \times \Gamma \times \mathcal{X}^\infty$ where the new parameter γ indexes the signal-generating distribution, and the parameters θ and γ jointly determine a family $(P_{\theta,\gamma})_{\theta\in\Theta,\gamma\in\Gamma}$ of conditional distributions over signals. The key distinction between θ and γ is that only θ is payoff-relevant. We'll use P^i to denote agent i's subjective prior belief on Ω, which is common knowledge to all agents.

If people do not in fact have dogmatic beliefs about the signal-generating distribution, a natural question is whether modeling agents in this way is still a good abstraction, in the sense that the qualitative insights of this model are robust to introduction of a small amount of model uncertainty. Acemoglu, Chernozhukov and Yildiz (2015) demonstrate one important sense in which this is not so.

7.1.3 Failure of Asymptotic Agreement

For any infinite sequence $\mathbf{x} \in \mathcal{X}^\infty$, write

$$\phi^i_{\theta,t} \equiv P^i(\theta \mid x_1, \ldots x_t)$$

for the posterior probability that agent i assigns to θ following the first t realizations of the sequence $\mathbf{x}$. Further define

$$\phi^i_{\theta,\infty}(\mathbf{x}) = \lim_{t\to\infty} \phi^i_{\theta,t}(\mathbf{x}) \tag{7.1}$$

to be the asymptotic posterior probability that agent i assigns to θ along sequence $\mathbf{x}$.

Definition 7.1. *Say that* asymptotic agreement *occurs if for each agent i,*

$$P^i(\phi^1_{\theta,\infty} = \phi^2_{\theta,\infty}) = 1 \quad \forall \theta \in \Theta$$

That is, both agents believe their asymptotic beliefs will be identical.

When agents hold a dogmatic belief about the signal-generating distribution, asymptotic agreement occurs whenever the parameter is identified (Proposition 18). But Acemoglu, Chernozhukov and Yildiz (2015) show that asymptotic agreement can fail when an arbitrarily small amount of model uncertainty is introduced. The basic idea behind this fragility can be seen through this following example from their paper.

Let $\Theta = \{A, B\}$, with each agent i's prior about the parameter denoted by $\pi^i \equiv (\pi^i_A, \pi^i_B)$. Agent i believes that signals are generated iid from the set $\{a, b\}$ with state-dependent distribution

	a	b
A	γ	$1-\gamma$
B	$1-\gamma$	γ

where γ is unknown and distributed according to G^i with density

$$g^i(\gamma) = \begin{cases} \varepsilon + \frac{1-\varepsilon}{\lambda} & \text{if } \gamma \in (\gamma^i - \lambda/2, \gamma^i + \lambda/2) \\ \varepsilon & \text{otherwise} \end{cases}$$

for some $\gamma^i > 1/2$. Assume that γ^1 and γ^2 are different from one another. This density is depicted in Figure 7.1.

The limit as $\varepsilon \to 0$ and $\lambda \to 0$ returns the model in which each agent i dogmatically believes the signal structure to be given by

	a	b
A	γ^i	$1-\gamma^i$
B	$1-\gamma^i$	γ^i

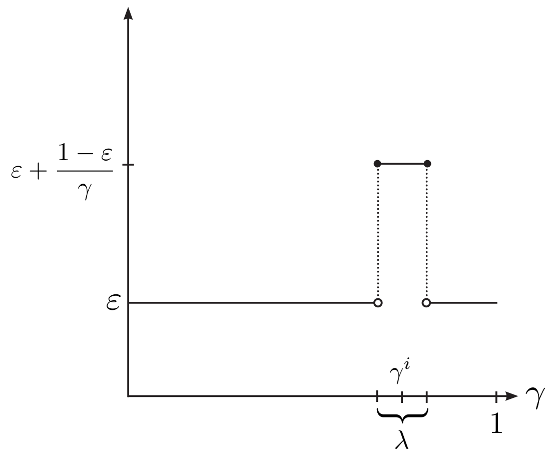

Figure 7.1: Depiction of g^i.

At this limit, asymptotic agreement holds.

Now suppose ε and λ are strictly positive and λ is small (specifically, let $\lambda < |\gamma^1 - \gamma^2|$ and suppose $\gamma^i - \frac{\lambda}{2} > \frac{1}{2}$ for each agent i). As in Section 6.2, define

$$n_t(\mathbf{x}) \equiv \#\{1 \le t' \le t : \mathbf{x}_{t'} = a\} \quad \forall \mathbf{x} \in \mathcal{X}^\infty$$

to be the count of a-realizations among the first t realizations of $\mathbf{x}$, and let

$$\rho(\mathbf{x}) = \lim_{t\to\infty} n_t(\mathbf{x})/t \quad \forall \mathbf{x} \in \mathcal{X}^\infty$$

be the asymptotic frequency of a-realizations along $\mathbf{x}$.

The following lemma provides a simple expression for the agent's asymptotic belief (7.1) on the set of sequences $\widetilde{\mathcal{X}}^\infty \subseteq \mathcal{X}^\infty$ where the limiting frequency $\rho(\mathbf{x})$ exists.

Lemma 3 (Acemoglu, Chernozhukov and Yildiz (2015)). *For every sequence* $\mathbf{x} \in \widetilde{\mathcal{X}}^\infty$,

$$\phi^i_{A,\infty}(\mathbf{x}) = \left(1 + \frac{1-\pi^i_A}{\pi^i_A} \cdot \frac{f^i_B(\rho(\mathbf{x}), 1-\rho(\mathbf{x}))}{f^i_A(\rho(\mathbf{x}), 1-\rho(\mathbf{x}))}\right)^{-1}$$

where $\frac{f^i_B(\rho(\mathbf{x}),1-\rho(\mathbf{x}))}{f^i_A(\rho(\mathbf{x}),1-\rho(\mathbf{x}))}$ *is the asymptotic likelihood ratio under agent i's subjective model.*

In the running example of this section, the asymptotic likelihood ratio can be simplified to

$$\frac{f^i_B(\rho, 1-\rho)}{f^i_A(\rho, 1-\rho)} = \frac{g^i(1-\rho)}{g^i(\rho)}$$

This ratio takes on either of three possible values. For any $\rho \in (\gamma^i - \lambda/2, \gamma^i + \lambda/2)$,

$$\frac{g^i(1-\rho)}{g^i(\rho)} = \frac{\varepsilon\lambda}{1-\varepsilon(1-\lambda)}$$

which converges to zero as ε and λ grow small (implying $\phi^i_{A,\infty} \to 1$). By a mirror argument, if the limiting frequency of a-realizations is some $\rho \in (1-\gamma^i-\lambda/2, 1-\gamma^i+\lambda/2)$, then

$$\frac{g^i(1-\rho)}{g^i(\rho)} = \frac{1-\varepsilon(1-\lambda)}{\varepsilon\lambda}$$

which converges to ∞ as ε and λ grow small (implying $\phi_{A,\infty} \to 0$). For all other limiting frequencies, the asymptotic likelihood ratio is simply $\frac{g^i(1-\rho)}{g^i(\rho)} = 1$. These unlikely signal sequences are considered possible but uninformative about the parameter.

Applying Lemma 3, Figure 7.2 depicts agent i's asymptotic posterior as a function of the limiting signal frequency.

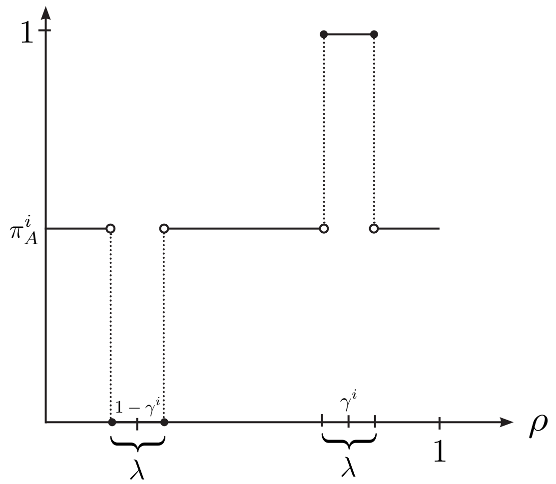

Figure 7.2: Agent i's asymptotic posterior in the limit as $\varepsilon \to 0$.

In the limit as $\varepsilon \to 0$ and $\lambda \to 0$, each agent i is increasingly sure that the limiting frequency ρ will either be close to γ^i or $1-\gamma^i$, so he believes that he will (approximately) learn the parameter. But when a sequence of signals has a long-run frequency that leads agent 1 to learn $\theta = A$ or $\theta = B$, agent 1 knows that this sequence has led agent 2 to consider the signal uninformative, in which case agent 2's limiting belief is the same as his prior. Likewise whenever agent 2 believes the signal sequence to be informative about θ, he knows that agent 1 considers the signal sequence to be uninformative. So not only does asymptotic agreement fail, but we have the stronger conclusion that the limiting beliefs ϕ^1_∞ and ϕ^2_∞ are different on *all* sample paths. Figure 7.3 depicts $|\phi^1_{A,\infty} - \phi^2_{A,\infty}|$ as a function of the limiting signal frequency.

To summarize, asymptotic agreement holds in the limiting model $\varepsilon = 0, \lambda = 0$ (with no model uncertainty), but fails when the model is perturbed to include an arbitrarily small amount of model uncertainty via $\varepsilon > 0, \lambda > 0$.

REMARK 7.1. As in Section 6, there is no ground truth—whether asymptotic agreement does or doesn't hold is determined solely with respect to the agents' subjective beliefs.

REMARK 7.2. In this example, the two agents' prior beliefs on $\Theta \times \Gamma$ are absolutely continuous with respect to one another. So Proposition 19 tells us that their beliefs about future signal realizations will eventually merge. But (θ, γ) is not identified: For example, $(A, 1)$ and $(B, 0)$ identically lead to a degenerate distribution on the infinite sequence of a-realizations. Thus asymptotic agreement about the expanded parameter (θ, γ) is not guaranteed from the results of Sections 6.3 and 6.4.

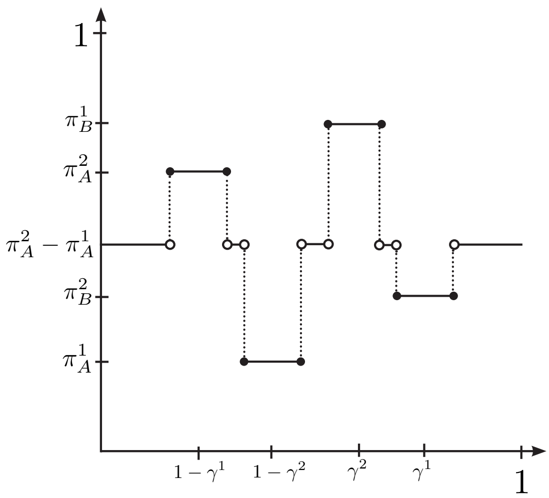

Figure 7.3: Asymptotic disagreement $|\phi^1_{A,\infty} - \phi^2_{A,\infty}|$ in the limit as $\varepsilon \to 0$, for parameter values $\pi_B^1 > \pi_A^2 > \pi_B^2 > \pi_A^1$.

7.2 Misspecified Learning

Next suppose the agent is not simply uncertain about the signal-generating distribution, but in fact rules out the true distribution.

EXAMPLE 7.1. Let $\Theta = \{A, B, C\}$ where the conditional distributions over signal realizations $\{a, b\}$ are given as follows:

	a	b
A	4/5	1/5
B	1/2	1/2
C	2/3	1/3

The agent has a uniform prior on $\{A, B\}$, but the true parameter is C. Given repeated independent observations from the distribution $(2/3, 1/3)$, will the agent's beliefs converge and if so to what limiting belief?

7.2.1 Role of KL Divergence

Intuitively, we may expect that the agent's beliefs converge to certainty of the parameter whose distribution is "closer" to the true distribution. The right notion of closeness here turns out to be KL Divergence (Section 5.1.2).

Here is a heuristic argument for how KL divergence emerges. Suppose the agent only considers parameter values $\theta = A$ and $\theta = B$ to be possible, where the prior probability of $\theta = A$ is $\pi \in (0, 1)$. We'll use $f_\theta(x)$ to denote the conditional probability of signal realization x when the parameter is θ. The agent observes a sequence of signals drawn iid according to f_{θ^*}, where the "true" parameter value θ^* may be different from both A and B.

For any signal sequence $\mathbf{x}_t = (x_1, \ldots, x_t)$, the conditional probability of A can be rewritten

$$
\begin{aligned}
&\mathbb{P}(\theta = A \mid \mathbf{x}_t) \\
&= \left(1 + \frac{1-\pi}{\pi}\left(\prod_{i=1}^{t} \frac{f_B(x_i)}{f_A(x_i)}\right)\right)^{-1} \\
&= \left(1 + \frac{1-\pi}{\pi}\left(\prod_{i=1}^{t} \frac{f_B(x_i)/f_{\theta^*}(x_i)}{f_A(x_i)/f_{\theta^*}(x_i)}\right)\right)^{-1} \\
&= \left(1 + \frac{1-\pi}{\pi}\exp\left(-\log\left(\prod_{i=1}^{t} \frac{f_{\theta^*}(x_i)}{f_B(x_i)}\right) + \log\left(\prod_{i=1}^{t} \frac{f_{\theta^*}(x_i)}{f_A(x_i)}\right)\right)\right)^{-1} \\
&= \left(1 + \frac{1-\pi}{\pi}\exp\left(-n \cdot \left(\frac{1}{t}\sum_{i=1}^{t} \log\left(\frac{f_{\theta^*}(x_i)}{f_B(x_i)}\right) - \frac{1}{t}\sum_{i=1}^{t} \log\left(\frac{f_{\theta^*}(x_i)}{f_A(x_i)}\right)\right)\right)\right)^{-1}
\end{aligned}
$$

and for large t this final display is approximately equal to

$$
\left(1 + \frac{1-\pi}{\pi}\exp\left(-t \cdot \left(D(f_{\theta^*} \| f_B) - D(f_{\theta^*} \| f_A)\right)\right)\right)^{-1} \tag{7.2}
$$

If $\theta^* \in \{A, B\}$, then either $D(f_{\theta^*} \| f_A) = 0 < D(f_{\theta^*} \| f_B)$ (in which case the expression in (7.2) converges to 1) or $D(f_{\theta^*} \| f_B) = 0 < D(f_{\theta^*} \| f_A)$ (in which case the expression in (7.2) converges to 0). In either case beliefs converge to certainty of the true parameter, as previously implied by Proposition 18 (Section 6.3).

Suppose now that $\theta^* \notin \{A, B\}$. Proposition 18 no longer applies: Doob (1949)'s consistency result is with respect to a P-measure 1 set of sequences, (where P is the agent's prior on $\Theta \times \mathcal{X}^\infty$), but in this example θ^* falls in the P-measure zero set on which consistency is not guaranteed. Indeed, in Section 6.3 we made no reference to a "true" distribution—consistency was demonstrated within the agent's subjective model.

But (7.2) is useful even when θ^* has zero probability under the agent's prior. Specifically, when $D(f_{\theta^*}\|f_A) < D(f_{\theta^*}\|f_B)$, then (7.2) converges to 1 as $t \to \infty$, yielding certainty of $\theta = A$, and when $D(f_{\theta^*}\|f_A) > D(f_{\theta^*}\|f_B)$, then (7.2) converges to zero as $t \to \infty$. So the agent's beliefs concentrate on the parameter that induces a distribution over signals that is closest in Kullback-Liebler divergence to the true distribution.

Berk (1966) establishes this result more generally. We'll use the notation of Section 6.1, introducing θ^* as new notation for the true parameter, and assuming that the observed signals are drawn iid according to the density f_{θ^*} (with P_{θ^*} denoting the induced measure on $\mathcal{X}^\infty$). To simplify exposition, assume that Θ is finite.

Proposition 23 (Berk (1966))**.** *Let*

$$A \equiv \arg\min_{\theta \in Supp(P)} D(f_{\theta^*}\|f_\theta)$$

be the set of parameters in the support of the agent's prior that minimize KL divergence to the true distribution. Then

$$\lim_{t\to\infty} P(A \mid X_1, \ldots, X_t) = 1 \quad P_{\theta^*}\text{-a.s.}$$

Example 7.2. Returning to Example 7.1, since

$$D(f_C\|f_A) = (2/3)\cdot\log\left(\frac{2/3}{4/5}\right) + (1/3)\cdot\log\left(\frac{1/3}{1/5}\right) \approx 0.021$$
$$D(f_C\|f_B) = (2/3)\cdot\log\left(\frac{2/3}{1/2}\right) + (1/3)\cdot\log\left(\frac{1/3}{1/2}\right) \approx 0.025$$

Proposition 23 implies that the agent's beliefs converge to certainty of $\theta = A$.

7.2.2 Berk Nash Equilibrium

Standard equilibrium concepts in game theory assume that players best-respond to correct and common beliefs. Esponda and Pouzo (2016) proposes a new equilibrium concept (modifying Nash equilibrium) that allows players to be misspecified. As this definition can be applied also within a single-agent setting, and as the notation is substantially lighter in this case, we start by defining Berk Nash equilibrium with one agent.

Single Agent Settings

There is a finite set of payoff-relevant states Ω, a finite set of signal realizations $\mathbb{S}$, and a finite set of actions $\mathbb{A}$. The agent holds a prior p over $\Omega \times \mathbb{S}$. Additionally, there is a finite set of consequences $\mathbb{Y}$, which are determined by the agent's action and the state via a feedback function $f : \mathbb{A} \times \Omega \to \mathbb{Y}$. The agent's payoff function is $u : \mathbb{A} \times \mathbb{Y} \to \mathbb{R}$.

The timing is as follows. First the agent chooses a strategy $\sigma : \mathbb{S} \to \Delta(\mathbb{A})$ mapping the observed signal into a distribution over actions. Then, the state and signal (ω, s) are drawn according to p, and the action $\sigma(s)$ is implemented. Finally, the consequence y is determined given the action and state (a, ω), and the agent obtains payoff $u(a, y)$.

There is an *objective* mapping $Q : \mathbb{S} \times \mathbb{A} \to \Delta(\mathbb{Y})$ from actions and signals into distributions over consequences, where

$$Q(y \mid s, a) = \sum_{\omega : f(\omega, a) = y} p(\omega \mid s) \quad \forall (y, s, a).$$

This is the conditional distribution over consequences that a Bayesian agent with knowledge of f, the action a, and the signal realization s would expect.

The agent does not know Q (or f). His *subjective model* $\mathcal{Q} = \langle \Theta, (Q_\theta)_{\theta \in \Theta} \rangle$ is a parametrized family of mappings $Q_\theta : \mathbb{S} \times \mathbb{A} \to \Delta(\mathbb{Y})$.

Definition 7.2. *The agent is* correctly specified *if there exists* $\theta \in \Theta$ *such that* $Q_\theta(\cdot \mid s, a) = Q(\cdot \mid s, a)$ *for all* $(s, a) \in \mathbb{S} \times \mathbb{A}$*; otherwise the agent is* misspecified.

The following example is adapted from Esponda and Pouzo (2016):

Example 7.3. A monopolist chooses a price a, which together with a random shock $\omega \sim \mathcal{N}(0, 1)$ determines demand

$$y = f(a, \omega) = \phi(a) + \omega.$$

The monopolist's payoff is $u(a, y) = a \cdot y$. Under the objective mapping f , the conditional distribution $Q(\cdot \mid a)$ is normal with mean $\phi(a)$ and variance 1. The monopolist's subjective model is instead the family $Q_\theta(\cdot \mid a)$ of normal distributions indexed to $\theta = (\theta_0, \theta_1) \in \mathbb{R} \times \mathbb{R}$, where each $Q_\theta(\cdot \mid a)$ is normal with mean $\theta_0 + \theta_1 a$ and variance 1, corresponding to a perceived feedback function

$$f_\theta(a, \omega) = \theta_0 + \theta_1 a.$$

If ϕ is not in fact affine in a, then the monopolist is misspecified. (This example did not include a signal.)

For any agent strategy $\sigma : \mathbb{S} \to \Delta(\mathbb{A})$, define

$$q_\sigma(s, a) \equiv p_S(s)\sigma(a \mid s)$$

to be the distribution on $\mathbb{S} \times \mathbb{A}$ induced by the strategy σ and the agent's prior p. Further define

$$K(\sigma, \theta) = \sum_{(s,a) \in \mathbb{S} \times \mathbb{A}} \left(\mathbb{E}_{Q(Y \mid s, a)} \left[\ln \frac{Q(Y \mid s, a)}{Q_\theta(Y \mid s, a)} \right] \right) q_\sigma(a, s)$$

to be the expected Kullback-Leibler divergence between $Q_\theta(\cdot \mid s, a)$ and the objective distribution $Q(\cdot \mid s, a)$, weighted by $q_\sigma \in \Delta(\mathbb{S} \times \mathbb{A})$.

Given the agent's strategy σ, the set of closest parameters (in weighted KL divergence) is

$$\Theta^*(\sigma) = \arg\min_{\theta \in \Theta} K(\sigma, \theta)$$

Definition 7.3. *A strategy profile σ is a* Berk-Nash equilibrium *if there exists a $\mu \in \Delta(\Theta)$ such that*

(a) $\mu \in \Delta(\Theta^*(\sigma))$; i.e., μ has support on the set of KL-minimizers.

(b) σ is optimal given μ; namely, $\sigma(a \mid s) > 0$ implies that

$$a \in \arg\max_{a' \in \mathbb{A}} \mathbb{E}_{\overline{Q}_\mu(y|s,a')}[u(a', y)]$$

where $\overline{Q}_\mu(y \mid s, a) = \int_\Theta Q_\theta(y \mid s, a)\mu(\theta)d\theta$ is the conditional distribution over consequences that is induced by μ.

Example 7.4. A researcher's project is either good or bad, $\Omega = \{g, b\}$. The researcher observes a reaction to the project, which is either positive or negative, $\mathbb{S} = \{+, -\}$ where (ω, s) are jointly distributed according to:

	$s = +$	$s = -$
$\omega = g$	$1/3$	$1/6$
$\omega = b$	$1/6$	$1/3$

The researcher observes the signal $s \in \mathbb{S}$ and decides whether to exert high or low effort towards developing the project, $A = \{H, L\}$. The unknown true quality of the project, and the researcher's effort, jointly determine a journal outcome in $\mathbb{Y} = \{A, R\}$ (accept or reject) according to the following function

$$f(a, \omega) = \begin{cases} A & (a, \omega) = (H, g) \\ R & otherwise \end{cases}$$

That is, the project is accepted if it is good and also the researcher's effort is high, and it is rejected otherwise. The researcher's payoff is

$$u(a, y) = \begin{cases} 1 & (a, y) = (H, A) \\ -1 & (a, y) = (H, R) \\ 2 & (a, y) = (L, A) \\ 0 & (a, y) = (L, R) \end{cases}$$

The true distribution $Q(y \mid a, s)$ is described by $Q(A \mid +, L) = Q(A \mid -, L) = 0$ (since the paper will not be accepted if effort is low) and

$$Q(A \mid +, H) = p(\{\omega : f(H, \omega) = A\} \mid +) = p(g \mid +) = 2/3$$
$$Q(A \mid -, H) = p(\{\omega : f(H, \omega) = A\} \mid -) = p(g \mid -) = 1/3$$

since conditional on high effort, the probability of acceptance is equal to the probability that the paper is good. These conditional distributions are summarized as follows:

	A	R
$(+, H)$	$2/3$	$1/3$
$(-, H)$	$1/3$	$2/3$
$(+, L)$	0	1
$(-, L)$	0	1

Suppose the researcher's subjective model allows only for the parameters θ_1 and θ_2 which are indexed to the following conditional distributions:

	A	R
$(+, H)$	$3/4$	$1/4$
$(-, H)$	$1/2$	$1/2$
$(+, L)$	0	1
$(-, L)$	0	1

	A	R
$(+, H)$	$2/3$	$1/3$
$(-, H)$	$1/3$	$2/3$
$(+, L)$	$1/10$	$9/10$
$(-, L)$	$1/10$	$9/10$

The distribution on the left, Q_{θ_1}, overestimates the value of hard work, and the distribution on the right, Q_{θ_2}, is overly optimistic about the probability of acceptance given low effort. Is the strategy profile $\sigma(+) = H$, $\sigma(-) = L$ (in which the research exerts high effort after a positive signal and low effort after a low signal) a Berk Nash equilibrium?

The distribution q_σ assigns probability $1/2$ to $(+, H)$ and to $(-, L)$. So

$$\begin{aligned} K(\sigma, \theta) = \frac{1}{2} \left(\sum_{y \in \{A,R\}} Q(y \mid +, H) \cdot \ln \left(\frac{Q(y \mid +, H)}{Q_\theta(y \mid +, H)} \right) \right) \\ + \frac{1}{2} \left(\sum_{y \in \{A,R\}} Q(y \mid -, L) \cdot \ln \left(\frac{Q(y \mid -, L)}{Q_\theta(y \mid -, L)} \right) \right) \end{aligned}$$

and thus

$$\begin{aligned} K(\sigma, \theta_1) &= \frac{1}{2} \cdot \left(\frac{2}{3} \ln \left(\frac{2/3}{3/4} \right) + \frac{1}{3} \ln \left(\frac{1/3}{1/4} \right) \right) \approx 0.0038 \\ K(\sigma, \theta_2) &= \frac{1}{2} \cdot \ln \left(\frac{1}{9/10} \right) \approx 0.02 \end{aligned}$$

Hence θ_1 is the unique minimizer of KL divergence to the true distribution, i.e., $\Theta^*(\sigma) = \{\theta_1\}$.

Only $\mu = \delta_{\theta_1}$ (a point mass at θ_1) satisfies Part (a) of Definition 7.3, and the distribution $\overline{Q}_\mu$ in Part (b) of Definition 7.3 simplifies to Q_{θ_1}. To determine whether σ is a Berk Nash equilibrium, it remains to verify that σ satisfies the optimality condition in Part (b) of Definition 7.3.

Suppose the signal realization is $s = +$. Then the action H yields an expected payoff of

$$\mathbb{E}_{Q_{\theta_1}(y|+,H)}[u(H,y)] = 1 \cdot \frac{3}{4} - 1 \cdot \frac{1}{4} = \frac{1}{2}$$

while the action L yields an expected payoff of

$$\mathbb{E}_{Q_{\theta_1}(y|+,L)}[u(L,y)] = 0$$

so $a = H$ is indeed optimal.

Suppose the signal realization is $s = -$. Then the action H yields an expected payoff of

$$\mathbb{E}_{Q_{\theta_1}(y|-,H)}[u(H,y)] = 1 \cdot \frac{1}{2} - 1 \cdot \frac{1}{2} = 0$$

while the action L yields an expected payoff of

$$\mathbb{E}_{Q_{\theta_1}(y|-,L)}[u(L,y)] = 0.$$

So $a = L$ is a best reply, and we conclude that σ is a Berk Nash equilibrium.

In sum, we have shown that the strategy σ is a best reply to a point mass on the unique parameter that minimizes KL divergence to the distribution over consequences induced by σ. In this sense the strategy σ is internally consistent with respect to the agent's misspecified model.

Exercise 7.1 (G). *Solve for all remaining pure-strategy Berk Nash equilibria in Example 7.4, or prove that there are none other.*

Simultaneous-Move Games

We turn now to the definition of Berk Nash equilibrium in simultaneous-move games. There is a set of players I, a set of payoff-relevant states Ω, a set of signal profiles $\mathbb{S} = \times_i \mathbb{S}_i$, and a probability distribution p over $\Omega \times \mathbb{S}$ whose marginals have full-support. There is a set of action profiles $\mathbb{A} = \times_i \mathbb{A}_i$, a set of *consequence* profiles $\mathbb{Y} = \times_i \mathbb{Y}_i$, and a profile of *feedback functions* $f = (f_i)_{i \in \mathcal{I}}$ where each $f_i : \mathbb{A} \times \Omega \to \mathbb{Y}_i$ maps outcomes in $\Omega \times \mathbb{A}$ into consequences for player i. Agents have payoff functions $u_i : \mathbb{A}_i \times \mathbb{Y}_i \to \mathbb{R}$.

The timing of the game is as follows: First, the state and signal (ω, s) are drawn according to p. Then each player i privately observes his own signal s_i and chooses an action a_i. The profile of consequences is determined via f as a function of the action profile and the state, and payoffs are realized.

For any player i, action $a_i \in \mathbb{A}_i$, and consequence $y_i \in \mathbb{Y}_i$, let

$$\Lambda^i(a_i, y_i) = \{(\omega, a_{-i}) : f_i(a_i, a_{-i}, \omega) = y_i\}$$

be the state and opponent action profiles that induce consequence y_i given player i's choice of a_i. The *objective distribution* over player i's consequences is $Q^i_\sigma : \mathbb{S}_i \times \mathbb{A}_i \to \Delta(\mathbb{Y}_i)$, where

$$Q^i_\sigma(y_i \mid s_i, a_i) = \sum_{(\omega, a_{-i}) \in \Lambda^i(a_i, y_i)} \sum_{s_{-i} \in S_{-i}} \left(\prod_{j \neq i} \sigma^j(a^j \mid s^j) \right) \cdot p_{\Omega \times S_{-i} \mid S_i}(\omega, s_{-i} \mid s_i)$$

for all $(s_i, a_i, y_i) \in \mathbb{S}_i \times \mathbb{A}_i \times \mathbb{Y}_i$. This is the conditional distribution over consequences that a Bayesian agent with knowledge of f, the strategy profile σ, and the signal realization s_i would expect.

The subjective model $\mathcal{Q} = \langle \Theta, (Q_\theta)_{\theta \in \Theta} \rangle$, with $\Theta = \prod_{i \in \mathcal{I}} \Theta^i$ and $Q_\theta = (Q^i_{\theta_i})_{i \in \mathcal{I}}$, describes the set of distributions over consequences that each player considers possible. Each player's parameter set Θ_i indexes distributions $Q^i_{\theta_i} : \mathbb{S}_i \times \mathbb{A}_i \to \Delta(\mathbb{Y}_i)$.

Definition 7.4. *A game is* correctly specified given σ *if for all players i, there exists $\theta_i \in \Theta_i$ such that $Q^i_{\theta_i}(\cdot \mid s_i, a_i) = Q^i_\sigma(\cdot \mid s_i, a_i)$ for all $(s_i, a_i) \in \mathbb{S}_i \times \mathbb{A}_i$; otherwise the game is* misspecified given σ. *A game is* correctly specified *if it is correctly specified for all σ; otherwise it is* misspecified.

For any strategy profile σ, define

$$q_{\sigma_i}(s_i, a_i) \equiv \sigma_i(a_i \mid s_i) p_{S_i}(s_i)$$

For any strategy profile σ, define

$$K_i(\sigma, \theta_i) = \sum_{(s_i, a_i) \in \mathbb{S}_i \times \mathbb{A}_i} \left(\mathbb{E}_{Q^i_\sigma(\cdot \mid s_i, a_i)} \left[\ln \frac{Q^i_\sigma(Y_i \mid s_i, a_i)}{Q^i_{\theta_i}(Y_i \mid s_i, a_i)} \right] \right) q_{\sigma_i}(s_i, a_i)$$

to be the expected Kullback-Leibler divergence between $Q_{\theta_i}(\cdot \mid s_i, a_i)$ and the objective distribution $Q^i_\sigma(\cdot \mid s_i, a_i)$, weighting (s_i, a_i) pairs according to $q_{\sigma_i}(s_i, a_i)$.

The set of closest parameters is

$$\Theta_i(\sigma) = \arg\min_{\theta_i \in \Theta_i} K_i(\sigma, \theta_i)$$

Definition 7.5. *A strategy profile σ is a* Berk-Nash equilibrium *if for all players i, there exists a $\mu_i \in \Delta(\Theta_i)$ such that*

(a) $\mu_i \in \Delta(\Theta_i(\sigma))$; i.e., μ has support on the set of KL-minimizers.

(b) σ_i is optimal given μ_i; namely, $\sigma_i(a_i \mid s_i) > 0$ implies that

$$a_i \in \arg\max_{\bar{a}_i \in \mathbb{A}_i} \mathbb{E}_{\overline{Q}^i_{\mu_i}(\cdot \mid s_i, \bar{a}_i)} [u_i(\bar{a}_i, Y_i)]$$

where $\overline{Q}^i_{\mu_i}(\cdot \mid s_i, \bar{a}_i) = \int_{\Theta_i} Q^i_{\theta_i}(\cdot \mid s_i, a_i) \mu_i(\theta_i) d\theta_i$ is the distribution over consequences of player i, conditional on $(s_i, a_i) \in \mathbb{S}_i \times \mathbb{A}_i$, induced by μ_i.

Remark 7.3. This definition is equivalent to Nash equilibrium when (a) is replaced with the condition that players have correct beliefs; i.e., $\overline{Q}^i_{\mu_i} = Q^i_\sigma$.

Proposition 24 (Esponda and Pouzo (2016))**.** *A Berk-Nash equilibrium exists.*

Building on Proposition 23, several authors have examined convergence of misspecified learning processes where—different from Berk (1966)'s setting—signals are endogenous to actions chosen by agents (Nyarko, 1991; Fudenberg, Romanyuk and Strack, 2017; Heidhues, Koszegi and Strack, 2021). The stable outcomes under many of these processes turn out to correspond to Berk Nash equilibria or a refinement of this set. Some recent works on this topic include Esponda and Pouzo (2016), Esponda, Pouzo and Yamamoto (2021), Bohren and Hauser (2021), Fudenberg, Lanzani and Strack (2020), Esponda, Pouzo and Yamamoto (2021) and Frick, Iijima and Ishii (2022).

7.3 Additional Exercises

Exercise 7.2 (G). *There are two states of the world, $\theta \in \{A, B\}$. A news source receives an infinite sequence of signals about this state of the world drawn iid according to the following signal structure*

	a	b
$\theta = A$	$3/4$	$1/4$
$\theta = B$	$1/4$	$3/4$

This news source is biased. When it observes the signal realization a, it reports a, but conditional on observing the signal realization b, it reports this b with probability $1 - \lambda$ and otherwise falsely reports a (where λ is constant across time). You are aware that the news source is biased and dogmatically believe that $\lambda = 1/2$.

Suppose the true state is $\theta = B$, and you observe the infinite sequence of news reports. Provide a condition (potentially empty) on the true value of λ such that your asymptotic belief is that the state is $\theta = A$. Interpret this result.