Oil and Gas on Indian Reservations: Statistical Methods Help to Establish Value for Royalty Purposes

Mary S. Fowler
Worcester State College

Joseph B. Kadane
Carnegie Mellon University

Journal of Statistics Education Volume 14, Number 3 (2006),
ww2.amstat.org/publications/jse/v14n3/kadane.html

Copyright © 2006 by Mary S. Fowler and Joseph B. Kadane all rights reserved. This text may be freely shared among individuals, but it may not be republished in any medium without express written consent from the authors and advance notification of the editor.

Key Words: Expectation; Law; Location-scale family

Abstract

Part of the history of oil and gas development on Indian reservations concerns potential underpayment of royalties due to under-valuation of production by oil companies. This paper discusses a model used by the Shoshone and Arapaho tribes in a lawsuit against the Federal government, claiming the Government failed to collect adequate royalties. Portions of the case have been settled out of court with compensation paid to the Tribes. Other portions remain pending. This material can be used as a real example in a calculus-based probability and statistics course.

1. Introduction

The Eastern Shoshone and Northern Arapaho tribes of the Wind River Reservation of Wyoming own land that contains deposits of natural gas and oil. The United States government regulates oil and gas leasing on Indian reservations and supervises the collection of royalties. Federal regulations include provisions which establish minimum values of oil and gas for royalty purposes. Some of these regulations were designed to combat actual or perceived manipulations or under-valuations by integrated oil companies, who have incentives to report low values to reduce their royalty payments. The Shoshone and Arapaho tribes of the Wind River Reservation sued the federal government for failing to collect adequate royalties from the oil companies in accordance with applicable regulations. The question in the case was whether the Government computed royalties correctly, and, if not, how much money was due to the tribes as a result. In the matter of the Shoshone and Arapaho Tribes of the Wind River Reservation v. United States (Docket Nos. 458-79-L and 459-79-L in the United States Court of Federal Claims), the second
author was an expert witness on behalf of the Tribes.

We develop a probability model that estimates the money due to the tribes for a given distribution of oil and gas prices. The case was settled before data could be used to estimate this distribution; hence, this example demonstrates the power of applying probability in a legal application. This problem offers two things to a calculus-based probability class; first, it is nice to have a non-textbook example of the relevance of probability both through direct calculus application and computer simulation and second, it introduces the class to the application of probability and statistics to the legal field.

In Section 2, we present the historical background of Indian reservations and US involvement in the sale of oil and gas from Indian land. In Section 3, we describe Major Portion Analysis, the federally regulated method to establish royalty payments for gas and oil. In Section 4, we give an example of Major Portion Analysis and present a probability model for the loss of revenues due to a failure to perform Major Portion Analysis. Section 5 uses the probability model to develop an expression for the expected value of the revenue lost by the Indians, in general and in the special case of the normal and t-distributions. In Section 6, the losses are estimated through a computer simulation for a distribution that has a form for loss of revenue that is not mathematically tractable. Section 7 gives some additional problems appropriate for a student in a calculus based probability class.

2. Indian Reservations and US Involvement

2.1 Historical Introduction

In the early 1800s the Eastern Shoshone were a nomadic people that ranged from Southwestern Montana to Southwestern Wyoming. For websites about Native American History, see http://digital.library.okstate.edu/KAPPLER/index.htm, http://www.hanksville.org/NAresources/indices/NAhistory.html, http://www.csulb.edu/projects/ais.

By the mid-1800s, there was a migration of settlers from the eastern US to and across tribal lands. The native tribes resisted and it soon became clear that there was a need to settle and clarify land rights and other issues between the US government and the tribal people. In 1863, the first treaty between the US government and the leaders of the Shoshone tribe was signed. [See Kappler (1904).]

In 1878, the federal government authorized a band of starving Northern Arapahos to winter on the Wind River Reservation against the wishes of the Shoshone tribe. The Northern Arapaho tribe remains there to this day. In 1937, the Shoshone tribe successfully sued the U.S. government in the Court of Claims of the United States for half the value of the reservation because the US had essentially given half the Reservation to the Arapaho in violation of the Shoshone treaties. The Shoshone received cash and the Shoshone and Arapaho tribes were confirmed as joint owners of the Wind River Reservation.

2.2 Natural Resource Issues

The federal government always has and continues to supervise many tribal activities, particularly involving reservation natural resources. Although the oil and gas on most tribal reservations is beneficially owned by the tribe, the federal government has established guidelines for the leasing of the land to oil and gas producers. Any oil or gas lease on a reservation must be approved by the federal government. The federal government collects money from the company leasing the land and transfers it to the tribe. The government has defined in the Code of Federal Regulations methods to be used to determine the value of the oil and gas produced from Indian reservations for royalty purposes.
In 1979, the Shoshone and Arapaho tribes of the Wind River Reservation sued the U. S. government in connection with the selling of oil and gas, claiming that the federal government had not followed the method prescribed by regulation for valuing oil and gas for royalty purposes. One of the questions in this case was whether and to what extent the tribes would have benefited had it done so. For material on the specific case, see [www.usdoj.gov/osg/briefs/2004/2pet/7pet/2004-0929.pet.app.pdf](http://www.usdoj.gov/osg/briefs/2004/2pet/7pet/2004-0929.pet.app.pdf).

### 3. Major Portion Analysis

The landowner and the oil company lessee both have interests in oil and gas revenues. The petroleum company owes the landowner a royalty payment, usually a percentage of the sale price, for each barrel of oil or each thousand cubic feet of natural gas sold and has the responsibility to market production competently, so that the landowner receives its fair share of the value of production through royalties. Because the petroleum industry has many vertically integrated companies that contain an entity that produces gas or oil and affiliated ones that process and market it, a producer may be in position to artificially depress the price of the product received by the production affiliate to decrease the royalty payment to the landowner, and recoup profit on resale by other affiliates. The federal government has deemed it necessary to have alternative methods of valuing oil and gas to address this potential abuse.

For at least 50 years, federal regulations, and Indian leases themselves, have provided that value for royalty purposes of oil and gas be determined using a “major portion analysis.” In a major portion analysis the value of the oil or gas is to be no less than the highest price “paid or offered” at the time of production for the “major portion” (50% plus one barrel of oil or one thousand cubic feet of natural gas) of production. The highest price “paid or offered” means the price at which the transaction was conducted, unless a higher price was offered and refused, in which case the highest price “paid or offered” is the highest price offered. The relevant transactions used to determine the price of the major portion include only transactions within a given month between non-affiliated entities, also known as arm-length transactions, of like quality oil or gas, from the same field or area. Product sold at less than the major portion price is to be valued at the major portion price; product sold at more than the major portion price is to be valued at the sale price, when calculating royalty payments to landowners. These rules are set forth in the Code of Federal Regulation, specifically at 30 C.F.R. section 206.52 (oil) and 30 C.F.R. sections 206.172 and 173 (gas) [www.gpoaccess.gov/cfr/index.html](http://www.gpoaccess.gov/cfr/index.html). Because the major portion value depends on other transactions going on throughout a given month involving multiple operators who typically do not share pricing information with each other, it must be calculated by the federal government well after the transaction occurs.

For much of the period in dispute, the federal government did not perform a major portion analysis on the oil and gas sold on behalf of the Shoshone and Arapaho tribes of the Wind River Reservation. Instead, royalty payments were based on actual sale price, including non-arm's length sales between affiliates. As a result, the question arose whether and to what extent the tribes would have benefited from increased royalties had it done so.

We can address this question in a two-stage process:

1. Explore the variables that affect the loss of royalties and how the total loss of royalties is related to these variables. We need to find a functional form that represents lost royalties as a function of the distribution of sale prices for gas and oil.
2. To investigate sales data to determine the distribution of sale prices and use this in the functions described in stage 1 to determine the actual loss of royalties.

In this article, we address only the question in stage 1. A settlement was achieved before the second stage of the analysis was reached. Carrying out a major portion analysis would have been controversial, as
quarrels were likely to surface about what oil is of “like quality” and what the “same field or area” means operationally. Despite these practical difficulties and ambiguities, some useful information can be gleaned from the mathematical development of royalty loss as a function of the distribution of sale prices.

4. A Model

Before describing the model, we review how to calculate the median of a weighted set of numbers. Suppose that 500 barrels of oil are sold at $14, 1000 barrels at $16 and 1000 barrels at $17, and that these were the best prices offered. Then the prices can be thought of as 500 14’s, followed by 1000 16’s and then 1000 17’s, so that the prices are ordered from lowest to highest. Then the midpoint, at which half the prices are less and half greater, is $16. However, it is a lot of effort to write 2500 numbers. Suppose, instead, that we think of the sales in terms of 500 barrel lots. Then the prices per 500 barrels of oil are 14, 16, 16, 17, and 17 (in order), and again the median is 16. With this as background, conducting a major portion analysis requires the following:

a. The larger of the best price offered and the price paid for each transaction is calculated (for “like quality” product in the “same field or area,” and including only “arms-length” transactions, called here the relevant transactions). These relevant transactions are weighted by the volume of product sold, and ordered by price. The median \( m \) of these best prices offered or paid is calculated. (The effect of the “plus one barrel of oil” and “plus 1000 cubic feet of gas” is trivial in the calculations that follow, and is omitted for simplicity).

b. For each barrel of oil, the larger of \( m \) and the best price offered or paid for that barrel, is calculated. This is the royalty value of the barrel of oil. Royalty is due on this amount.

To continue the numerical example, (in lots of 500 barrels of oil), the royalty values would be 16, 16, 16, 17 and 17, respectively. Thus the effect of major portion analysis in this example is to increase the value of the 500 barrels sold at $14 to $16. This change in turn would entitle the tribes to a greater royalty payment.

A typical royalty rate at the time was 1/6 the value of the oil. In this hypothetical example, the 500 barrels sold at $14 had a royalty value of $16. Hence the Tribes should have been paid \((16 - 14)(500)/6 = 166.67\) more than they were for this month. Over many years, and with interest due from the time the money should have been paid to the present, this can add up to a substantial sum of money owed to the Tribes.

Necessarily, the application of (b) cannot lower the royalty value to which the Tribes are entitled. Suppose that the best prices paid or offered for relevant transactions for a given field in a given month are \( X_1, \ldots, X_n \). Necessarily \( \min_{1 \leq i \leq n} X_i \leq \text{med} X_i \). When and only when this inequality is strict, the Tribes gain by the application of major portion analysis. How much can this be expected to amount to? This depends on the distribution of the \( X \)’s, as the analysis below shows.

Let \( X \) be a random variable whose cumulative distribution function (cdf) is \( F \). Also let

\[
Y = \max \left\{ X, F^{-1}(0.5) \right\}.
\]

Then \( Y \) has the distribution of the royalty value of the oil or gas, and \( Y - X \) is the gain in royalty value due to the performance of major portion analysis. The next section studies the expectation of \( Y - X \).
5. Analysis

Suppose $X$ has density $f(x)$, median $m^*$, and scale $s$. Then the integral from $m^*$ to infinity of $y - x$ is zero because $y = x$ in this range. Therefore

$$ E(Y - X) = \int_{m^*}^{\infty} (y - x) f(x) dx. \quad (2) $$

Let $t = (x - m^*)/s$. Then

$$ E(Y - X) = \int_{-\infty}^{0} \left[ m^* - (st + m^*) \right] f(st + m^*) dt = -s \int_{-\infty}^{0} t f(st + m^*) dt. \quad (3) $$

In a location-scale family, if $Y$ has a density $f(y)$ and $X = m^* + sY$, where $m^*$ and $s > 0$ are chosen numbers, then $X$ has density $sf(sy + m^*)$. In such a family, the density satisfies

$$ sf(st + m^*) dt = g(t) dt \quad (4) $$

independent of $m^*$ and $s$. Thus $g(t)$ is the density of that member of the location-scale family with $m^* = 0$ and $s = 1$.

If $X$ is in such a family,

$$ E(Y - X) = -s \int_{-\infty}^{0} t g(t) dt. \quad (5) $$

This representation shows that the expected gain to the tribes from major portion analysis depends critically on the scale parameter $s$. The advantage to the tribes of major portion analysis is proportional to $s$, with the constant of proportionality dependent on the shape of the distribution. Next, we evaluate the constant of proportionality for some examples.

**Example 1 Normal family.**

Suppose $X \sim N(m, s^2)$. Then $X$ is a member of the location-scale family of normal distributions. Hence equation (5) applies. The integral in (5) is evaluated as follows:

$$ I_1 = \int_{-\infty}^{0} t g(t) dt = -\int_{0}^{\infty} t g(t) dt = -\int_{0}^{\infty} t \frac{1}{\sqrt{2\pi}} e^{-t^2/2} dt. $$

Let $y = t^2/2$, so $dy = t dt$.

$$ I_1 = -\int_{0}^{\infty} \frac{e^{-y}}{\sqrt{2\pi}} dy = -\frac{1}{\sqrt{2\pi}}. $$
Then

$$E(Y - X) = s/\sqrt{2\pi}.$$  \hspace{1cm} (6)

For example, an average of $15 per barrel and a standard deviation of $2 per barrel, would lead, by formula (6), to a loss of royalty value of $\frac{2}{\sqrt{2\pi}} = .80 per barrel. If the oil produced on Tribal Lands were 100,000 barrels in a year, the loss in royalty value would have been$80,000. At a royalty rate of 1/6, this would have come to $80,000/6 =$13,333 for that year.

**Example 2** Now suppose $X$ has a $t$-distribution with $\nu$ degrees of freedom.

The $t$-distribution most readers may be familiar with is a standard $t$-distribution which has median 0 and variance $\nu/(\nu - 2)$ where $\nu$ is the degrees of freedom (see DeGroot and Schervish (2002), p. 407). (For degrees of freedom less than 2, the variance of the $t$-distribution does not exist). However, the $t$-distribution can be extended to be a location-scale family by allowing a linear function of a standard $t$, as follows:

Let

$$X = sW + m*$$

where $W$ has a standard $t$-distribution with $\nu$ degrees of freedom. Then $X$ has median $m*$, and variance $s^2\nu/(\nu - 2)$. (See DeGroot (1970, 2004), p. 42) for more on the location-scale extension of $t$-distributions).

While the $t$-distribution is defined for all positive $\nu$, it has a mean only if $\nu > 1$. Hence we restrict $\nu$ to that domain for this calculation. Again, $X$ is in a location-scale family, so again (5) applies, and the integral is

$$I_2 = -\int_0^\infty \frac{t}{\sqrt{\nu B\left(\frac{1}{2}, \frac{\nu}{2}\right)}} \left(1 + \frac{t^2}{\nu}\right)^{-\frac{\nu+1}{2}} \; dt.$$  

Where $B(a,b)$ is the beta function defined by $B(a,b) = \int_0^1 x^{a-1}(1-x)^{b-1} \; dx$ for $a > 0, b > 0$.

Let $c = \sqrt{\nu B\left(\frac{1}{2}, \frac{\nu}{2}\right)}$, and $w = 1 + \frac{t^2}{\nu}$, so $dw = 2t/\nu \; dt$.

Then
Then

\[ I_2 = \frac{-c \nu}{2} \int_{1}^{\infty} w^{-\frac{1}{2}(\nu+1)} \, dw = \frac{-c \nu}{2} \left( \frac{1-\nu}{2} \right)^{\frac{1}{\nu-1}} \]

\[ = \frac{-c \nu}{1-\nu} \left[ \frac{1}{\nu-1} \right] = \frac{c \nu}{\nu-1}. \]

The star for the normal distribution is displayed on the extreme right of Figure 1, because a normal distribution is the same as a \( t \)-distribution with infinite degrees of freedom.

Figure 1 displays the expected gain from major portion analysis in royalty valuation, \( E(Y - X) \), for \( t \)-distribution (as a function of the degrees of freedom, \( \nu \)), and for the normal distribution. Both calculations take the scale parameter \( s \) equal to one. The star for the normal distribution is displayed on the extreme right of Figure 1, because a normal distribution is the same as a \( t \)-distribution with infinite degrees of freedom.
That the location parameter \( m^* \) is irrelevant to \( E(Y - X) \) should not be a surprise, since adding a constant to \( X \) adds a constant to \( Y \), and hence leaves \( Y - X \) unchanged. Similarly it should not be a surprise that as the scale \( s \) increases, the expected gain to the tribes increases and conversely as \( s \) decreases so does the expected gain. As a limiting case, we note that as \( s \) approaches zero, all prices approach \( m^* \) and hence the gains to the tribes would go to zero. In every location-scale family with a mean, the expected gain in royalty value due to the tribes by conducting a major portion analysis is a constant times the scale. Like exact derivations using calculus, simulations are important tools for modern statisticians.

6. Simulations

In the previous section, we derive expressions for the expected loss, \( E(Y - X) \) in the special cases when \( X \) has a normal distribution and a \( t \)-distribution. However, the integral in equation (5) is not tractable for many distributions. Suppose, for example, the best prices paid or offered are believed to come from a gamma distribution, then \( E(Y - X) \) is not analytically tractable. We use \texttt{R}, a statistical language available as free software at [www.r-project.org](http://www.r-project.org). Students can download \texttt{R} and work through the following simulation. This exercise demonstrates the power of simulation without any major challenges for the student.

We can approximate \( E(Y - X) \) by simulating the process of major portion analysis. The following \texttt{R}-code achieves such a simulation:

```r
n = 100
dx = rgamma (n, shape = 15, scale = 1)
#this draws n=100 independent observations from a gamma
#distribution with alpha = shape = 15 and scale = beta = 1.
#The mean is alpha/beta = 15 and the variance is alpha/beta^2 = 15
m = median(x).
y=pmax (x,m)
# this yields a vector of length n whose ith
# elements is the larger of x[i] and m.
mean(y-x)
# this is the estimated average amount per observation
# of underestimation of the royalty value of the oil by
#virtue of not conducting major portion analysis.
sd(y-x)
# this computes the standard deviation of (y-x).
```

We did this and obtained: \( \text{mean}(y - x) = 0.942 \) and \( \text{sd}(y - x) = 1.29 \). If the simulation is repeated, different values for the mean and standard deviation will be obtained, because the random numbers drawn will differ.

The observations simulated from the gamma distribution are independent and identically distributed, and so we can apply the central limit theorem, which tells us that the standard deviation of the mean of \( Y - X \) is estimated by \( \text{sd}(Y - X)/\sqrt{n} \). Hence the larger our \( n \) is in our simulation, the more stable our estimate of the mean of \( Y - X \) will be. To see this more clearly we repeated the above simulation for \( n \) equal 10, 100, 1000, 10000, 100000 and obtained the following results.
Table 1: Simulation results from a gamma (15,1) distribution for various different sample sizes, \( n \).

| \( n \)  | mean(\( Y - X \)) | \( sd(Y - X) \) | \( sd(Y - X)/\sqrt{n} \) |
|-------|------------------|-----------------|-------------------|
| 10    | 1.45             | 2.35            | 0.743             |
| 100   | 1.19             | 1.82            | 0.182             |
| 1000  | 1.35             | 1.93            | 0.061             |
| 10000 | 1.35             | 1.89            | 0.0189            |
| 100000| 1.37             | 1.92            | 0.00607           |

The results reported in Table 1 are plotted in Figure 2. Each line represents a 2-standard deviation interval around the mean. As \( n \) increases from 10 to 100,000, i.e., as \( \log(n) \) increases from 1 to 5, uncertainty is reduced, as is predicted by the central limit theorem.

Figure 2: 95% confidence intervals for \( Y - X \) as a function of sample size for Gamma (15,1).
The above R-code can be used to simulate a major portion analysis from other distributions by substituting another random number generator for 

\[ x = \text{rgamma}(n, \text{shape} = 15, \text{scale} = 1). \]

For example, to simulate Example 1 numerically, where \( X \) has a normal distribution, we replace the above line with 

\[ x = \text{rnorm}(n, 15, 1). \]

Here the mean is equal to 15 and the variance is equal to 1. According to equation 6, the exact value of

\[ E(Y - X) = \frac{1}{\sqrt{2\pi}} = 0.3989. \]

We ran the simulation in R and obtained \( \text{mean}(Y - X) = 0.3738 \). If you try the same simulation, you will draw a different set of random numbers and hence your value for mean \( (Y - X) \) will differ.

Again, if you wish to simulate from a beta distribution with mean \( a/(a+b) = 1/3 \) and variance \( ab/(a+b)(a+b+1) = 1/18 \), this implies \( a = 1 \) and \( b = 2 \). Then the R-code would be

\[ x = \text{rbeta}(n, 1,2). \]

7. Some Additional Problems

The following results can be shown using calculus based probability theory and are good problems for probability students to solve on their own.

\begin{enumerate}
  \item Suppose that \( X \) is symmetric around its mean \( m \), and has variance \( \sigma_x^2 \). Then
    \begin{enumerate}
      \item \( \text{Var}(Y) = \text{Var}(X - Y) = \frac{1}{2} \sigma_x^2 - [E(X - Y)]^2. \)
      \item \( \text{Cov}(X, Y) = \frac{1}{2} \sigma_x^2. \)
      \item As a consequence, show that the correlation between \( X \) and \( Y \) is given by 
        \[ \rho(X, Y) = \sigma_x \sqrt{\frac{2 \sigma_y}{\sigma_x^2}} \],
        where \( \sigma_y \) is the standard deviation of \( Y \).
    \end{enumerate}
  \item Since the normal distribution is the limiting case of the \( t \)-distribution as \( \nu \to \infty \), it is natural to conjecture that the constant multiplying the scale factor in equation (6),
        \[ E(Y - X) = \frac{1}{\sqrt{2\pi}} = 0.3989, \]
        is the limit as \( \nu \to \infty \) of the constant multiplying the scale factor in equation (7),
        \[ \frac{\sqrt{\nu}}{(\nu - 1)E\left(\frac{1}{2\cdot \frac{\nu}{2}}\right)}. \]
        Prove that this is true.
  \item Show that the moment generating function of \( Y - X \) when \( X \sim N(m, \sigma_x^2) \) is
b. Use derivatives of this result to check the mean and second moment of \( Y - X \).

8. Conclusion

We have shown that the only way the tribes could achieve the benefits of a major portion analysis, without having one performed, would be if there were no variation below the median in best price offered or paid for the relevant transactions. Since this typically is not the case, we are led to the conclusion that the tribes lost revenue when the government has not performed a major portion analysis.

The case was settled after the presentation of the above model. Documents from the case are available on line through the PACER system. You can register for PACER at pacer.psc.uscourts.gov/register.html. Once registered, go to pacer.psc.uscourts.gov/pasco/cgi-bin/links.pl and choose the United States Federal Claims Court website: ecf.cofc.uscourts.gov/ and enter case number 79-458. We hope to demonstrate to students the power of probability modeling, both through calculus based theory and computer simulation and that student will begin to consider the field of legal statistics and a source of interesting applications with importance to our society.

Acknowledgment

The authors thank a class at the Department of Statistics, Pontifical Catholic University, Santiago, Chile for their helpful comments on an earlier draft. We also thank Richard Berley of Zion, Chestnut, Varnell, Berley and Slonim for his helpful comments.

References

DeGroot, M. H. and Schervish, M. J. (2002), *Probability and Statistics*, 3rd edition, Boston: Addison-Wesley.

DeGroot, M. H. (1970, 2004), *Optimal Statistical Decisions*, Hoboken, NJ: J. Wiley and Sons.

Kappler, C. J. (editor) (1904), *Indian Affairs: Laws and Treaties, Vol. II, Treaties*. Compiled and edited by Charles J. Kappler. Washington: Government Printing Office, 1904.

Mary S. Fowler
Department of Mathematics
Worcester State College
Worcester, MA 01602
U.S.A.

mfowler@worcester.edu

Joseph B. Kadane

Mary S. Fowler
Department of Mathematics
Worcester State College
Worcester, MA 01602
U.S.A.

mfowler@worcester.edu

Joseph B. Kadane
