Posterior Momemts of the Cauchy Distribution

David R. Wolf - wolf@lanl.gov
Explosives Technology and Applications
M4 Image Analysis Section, MS P940, LANL, 87545
1 Apr 1993

1. INTRODUCTION

The Cauchy distribution [1]-[2] is given by the probability density function for the observation, on a line in some plane, of particles that are radiated randomly from an omnidirectional radiator at some position in the plane. If the line is the x axis and the radiator is at position (x_0, y_0) then this distribution is given by $P(x \mid x_0, y_0) = \frac{y_0}{\pi\left(y_0^2 + (x - x_0)^2\right)}$.

Given N observations (receptions of particles) $\mathbf{x} = (x_1, \ldots, x_N)$ and given that the prior for the true position of the radiator is given by $P(x_0, y_0)$ then Bayes theorem states that the probability density function for the true position of the radiator is proportional to the product of the Cauchy distribution and the prior, $P((x_0, y_0) \mid \mathbf{x}) \propto P(\mathbf{x} \mid x_0, y_0)\, P(x_0, y_0)$. Since the observations are independent, $P(\mathbf{x} \mid x_0, y_0) = \prod_{i=1}^{N} P(x_i \mid x_0, y_0)$.

The use of the first posterior moment as an estimator for position is motivated by the fact that it is the minimum mean squared error (mmse) estimate for the mean of the distribution. It is straightforward to show that the n^{th} posterior moment is the mmse estimator for the n^{th} power of position. In particular, the combination of posterior averages $E[x^2 \mid \mathbf{x}] - E[x \mid \mathbf{x}]^2$ is the mmse estimator for the variance of the distribution. Thus we are interested in computing these moments.

In the rest of this brief paper we show how to compute these moments in closed form. To begin, the moments are computed when the prior is uniform on the line, and we consider moments of x (moments of y are similar). Later we discuss the computation of moments when the prior is uniform on any interval. The finite interval solution is discussed in the limit of a large interval and explicit agreement with the first solution is shown. An example is given which demonstrates the failure mode in the context of Sec. 3 in finding the first moment of the one-event posterior and the role of the Cauchy principle value as an estimator. The first moments of the two and three event posteriors are stated and contrasted. The estimator for the variance in the 2 observation case is given. We discuss forms for priors that allow closed form solutions for the moments to be found. We discuss an approximation to non-integer moments. Finally, we show how to compute the moment generating function in closed form.

2. INFINITE INTEGRATION INTERVAL

When the interval of integration is $(-\infty,\infty)$ then the problem of finding moments may be viewed as a contour integration problem. The zeros in the denominator of the Cauchy distribution are given by $x_i^{\pm} = x_i \pm iy$. Thus the integral of interest is proportional to

$$\int_{-\infty}^{\infty} x^k \Pi_{i=1}^{N} \frac{1}{y^2 + (x-x_i)^2}\, dx = \oint_C x^k \Pi_{i=1}^{N} \frac{1}{y^2 + (x-x_i)^2}\, dx = 2\pi i \Sigma_{i=1}^{N} R_i$$

where $R_i = \frac{(x_{i+})^k}{x_{i+} - x_{i-}} \Pi_{k \neq i} \frac{1}{(x_{i+} - x_{k+})(x_{i+} - x_{k-})}$ and integer k, $0 \le k < 2N - 1$.

3. FINITE INTEGRATION INTERVAL

When the interval of integration is finite the residue theorem does not apply. However, it is possible to expand the denominator to find

$$\Pi_{i=1}^{N} \frac{1}{y^2 + (x-x_i)^2} = \Sigma_{i=1}^{N} \left[\frac{C_{i+}}{x - x_{i+}} + \frac{C_{i-}}{x - x_{i-}} \right]$$

where $C_i^{\pm} = \frac{\pm 1}{x_{i+} - x_{i-}} \Pi_{k \neq i} \frac{1}{(x_i^{\pm} - x_{k+})(x_i^{\pm} - x_{k-})}$. Now proceed with the integration as usual, noting that[1] $\int \frac{x^k}{x-z}\,dx = \sum_{i=1}^{k} \frac{x^i z^{k-i}}{i} + z^k \mathrm{Log}(x-z) + c$. The result is that the desired posterior average is proportional to

$$\int_a^b x^k \Pi_{i=1}^{N} \frac{1}{y_i^2 + (x-x_i)^2}\,dx = \sum_{i=1}^{N} \Big(C_{i+} \Big(\sum_{m=1}^{k} \frac{1}{m} (x-x_{i+})^m (x_{i+})^{k-m} + (x_{i+})^k \mathrm{Log}(x-x_{i+}) \Big)$$

$$+ C_{i-} \Big(\sum_{m=1}^{k} \frac{1}{m} (x-x_{i-})^m (x_{i-})^{k-m} + (x_{i-})^k \mathrm{Log}(x-x_{i-}) \Big) \Big) \Big|_a^b ,$$

which simplifies to

$$\int_a^b x^k \Pi_{i=1}^{N} \frac{1}{y_i^2 + (x-x_i)^2}\,dx = \sum_{i=1}^{N} \Big(C_{i+} (x_{i+})^k \mathrm{Log}(x-x_{i+}) + C_{i-} (x_{i-})^k \mathrm{Log}(x-x_{i-}) \Big) \Big|_a^b ,$$

for[2] $0 \le k < 2N-1$.

4. LARGE INTEGRATION INTERVAL

When the interval of integration is large we expect the result of the last section, Sec. 3, to approach the result of Sec. 2. To show this, define $r_i^{\pm}$, $\theta_i^{\pm}$ by $r_i^{\pm} e^{i\theta_i^{\pm}} = x - x_i^{\pm}$. Then with this notation the result of Sec. 3 is

$$\int_a^b x^k \Pi_{i=1}^{N} \frac{1}{y_i^2 + (x-x_i)^2}\,dx = \sum_{i=1}^{N} \Big(C_{i+} (x_{i+})^k \big(\mathrm{Log}(r_{i+}) + i\theta_{i+} \big) + C_{i-} (x_{i-})^k \big(\mathrm{Log}(r_{i-}) + i\theta_{i-} \big) \Big) \Big|_a^b$$

letting $b \to \infty$, $a \to -\infty$[3] and using the result $\sum_{i=1}^{N} \big(C_{i+} (x_{i+})^k + C_{i-} (x_{i-})^k \big) = 0$[2] gives the large interval result $\int_a^b x^k \Pi_{i=1}^{N} \frac{1}{y_i^2 + (x-x_i)^2}\,dx = 2i\pi \,\mathrm{Im}\Big[\sum_{i=1}^{N} C_{i+} (x_{i+})^k \Big]$. This is immediately seen[2] to be the result of Sec. 2.

5. EXAMPLES

Here we give three examples which i) show how a moment may fail to exist in the context of Sec. 3, ii) give the first posterior moments for the 2 and 3 observation cases, and iii) give the estimator for the variance in the 2 observation case. Up to this point we have neglected the needed proportionality (normalization) constant. Taking this into account, the desired expectation with uniform prior on $[a,b]$ is given by

$$E[x^k \mid \mathbf{x}] = \left(\int_a^b x^k \prod_{i=1}^{N} \frac{y}{y^2 + (x-x_i)^2}\, dx\right) \Big/ \left(\int_a^b \prod_{i=1}^{N} \frac{y}{y^2 + (x-x_i)^2}\, dx\right).$$

In what follows we consider this expectation with $a \to -\infty$, $b \to \infty$. A simple Mathematica program for finding these moments is given in the appendix.

<u>Example 1, first moment of the 1 observation posterior</u>.

Here $k = 1 = 2N-1$ so we expect that the moment will not exist, since $k < 2N-1$ was required but is not satisfied (see Sec. 4). Clearly, the numerator term for $E[x \mid \mathbf{x}]$ with $a \to -\infty$, $b \to \infty$ does not exist because the integrand of the numerator is logarithmically divergent for $k = 1$. However it is still interesting to see explicitly how this integral is ill defined in the context of the results of Sec. 3. The expansion of the integrand is given by $C_{1+}x/(x - x_{1+}) + C_{1-}x/(x - x_{1-})$, and this integrates to $(x_{1+}\mathrm{Log}(x - x_{1+}) - x_{1-}\mathrm{Log}(x - x_{1-}))/2iy$. Converting to polar form gives $x_{1+}\mathrm{Log}(r_{1+}) - x_{1-}\mathrm{Log}(r_{1-}) + x_{1+}i\theta_{1+} - x_{1-}i\theta_{1-}$. As can be immediately seen, the *only* case when cancellation of the logarithms is to be expected for $a \to -\infty$, $b \to \infty$ independently is when $y = 0$. Thus the integral is ill defined. We may take the principle value of the integral as the object of interest. In that case, or when $y = 0$, the average is simply x_1.

<u>Example 2, first moments of the 2 and 3 observation posteriors</u>.

For two observations, the result is trivially given by $E[x \mid (x_1, x_2)] = (x_1 + x_2)/2$.

However, the pattern set (the average) does not continue; for 3 observations the first posterior moment is given by

$$E[x \mid x_1, x_2, x_3] = \frac{8y^2(x_1+x_2+x_3) - x_1^2(x_2+x_3) - x_2^2(x_1+x_3) - x_3^2(x_1+x_2) + 6x_1x_2x_3}{2(12y^2 - x_1^2 - x_2^2 - x_3^2 + x_1x_2 + x_2x_3 + x_3x_1)}$$

Interestingly, when $y = 0$ this is not the average. Another interesting observation is that the difference of this and the average of the three points has a numerator given by the symmetric form $(-2x_1 + x_2 + x_3)(x_1 - 2x_2 + x_3)(x_1 + x_2 - 2x_3)$. This pattern, too, does not continue. For example, for $N = 2$ the difference is zero.

Example 3, estimator for the standard deviation with two observations.

The second moment of the two observation posterior is given by $y^2 + (x_1^2 + x_2^2)/2$. This, along with the first moment result above gives $E[(x - E[x \mid \mathbf{x}])^2 \mid \mathbf{x}] = y^2 + (x_1 - x_2)^2/4$.

6. PRIORS

Briefly, for an infinite interval of integration, note that when the prior is any finite sum of the form $\Sigma_{i=0}^{n} a_i x^i$ then moments up to order $2(N-1) - n$ exist if this quantity is non-negative. For a finite interval of integration then any series which when multiplied by the likelihood is properly integrated term-by-term may be used as a prior. Rational functions where the numerator degree minus the denominator degree is less than or equal to $2(N-1)$ may be used as priors.

7. NON-INTEGER MOMENTS

For a finite interval of integration and uniform prior, the binomial series[4] may be used to develop an approximation in the following way. Let q be a non-integer real, and consider one of the integrals to be done for the moment: $\int_a^b x^q \Pi_{i=1}^{N}(\ldots)dx$. Make the expansion of the denominator of $(\ldots)$ as before, and in each of the terms make the substitutions $u = x - x_i^{\pm}$ respectively, with the + or - as appropriate for the term. Now, expand each of the resulting numerators $(u + x_i^{\pm})^q$ using the binomial series appropriate for the limit under consideration (depending upon whether $|u|x > \pm i$ or $|u|x < \pm i$ at the limits $x = a, b$ as appropriate). The series representing the integration may be truncated at whatever order is needed for the desired accuracy when it converges.

8. MOMENT GENERATING FUNCTION

The kth moment is given in terms of the kth derivative of the moment generating function by $E[x^k \mid \mathbf{x}] = \partial_\alpha^k E[e^{\alpha x} \mid \mathbf{x}] \big|_{\alpha = 0}$. The moment generating function for the N observation Cauchy distribution may be found in closed form. We have

$$E[e^{\alpha x} \mid \mathbf{x}] = \left(\int_a^b e^{\alpha x} \Pi_{i=1}^N \frac{1}{y^2 + (x - x_i)^2} dx\right) \Big/ \left(\int_a^b \Pi_{i=1}^N \frac{1}{y^2 + (x - x_i)^2} dx\right).$$

Expand the denominator in the numerator integral as in Sec. 3 to find

$$\int_a^b e^{\alpha x} \Pi_{i=1}^N \frac{1}{y^2 + (x - x_i)^2} dx = \int_a^b \Sigma_{i=1}^N \left(\frac{C_{i+} e^{\alpha x}}{x - x_{i+}} + \frac{C_{i-} e^{\alpha x}}{x - x_{i-}}\right) dx.$$

After making appropriate variable changes each of the terms in the sum on the right is of the form[5] [2]

$$Ce^{\alpha x_0} \int_{\alpha(a - x_0)}^{\alpha(b - x_0)} \frac{e^u}{u} du = Ce^{\alpha x_0} [Ei(\alpha(b - x_0)) - Ei(\alpha(a - x_0))].$$

Making the substitutions for the moment generating function yields

$$E[e^{\alpha x} \mid \mathbf{x}] = \frac{\Sigma_{i=1}^N Re(C_{i+} e^{\alpha x_{i+}} [Ei(\alpha(b - x_{i+})) - Ei(\alpha(a - x_{i+}))])}{\Sigma_{i=1}^N Re(C_{i+} [Log(b - x_{i+}) - Log(a - x_{i+})])}.$$

Clearly, all moments are easily accessible since $\partial_x Ei(x) = e^x / x$. As a check of the first moment[5], $\lim_{\alpha \to 0} [Ei(\alpha p) - Ei(\alpha q)] = Log(p) - Log(q)$.

ACKNOWLEDGMENTS

I would like to thank Kenneth Hanson for suggesting this problem. This work was supported by the United States Department of Energy under contract number W-7405-ENG-36.

REFERENCES

[1] K. M. Hanson, "Introduction to Bayesian Image Analysis", To appear in Proc. SPIE, vol. 1898, Medical Imaging: Image Processing, ed. M.H. Loew.

[2] Stephen F. Gull, "Bayesian Inductive Inference and Maximum Entropy", in Maximum Entropy and Bayesian Methods in Science and Engineering, G.J. Erickson and C. R. Smith (eds.), pp. 53-74, 1988, Kluwer Academic Publishers.

[3] M. Abramowitz and I. A. Stegun, Handbook of Mathematical Functions , (Dover, New York 1972).

FOOTNOTES

1 This follows immediately from the representation $\frac{x^k}{x-z} = \sum_{i=0}^{k-1} x^i z^{k-i-1} + \frac{z^k}{x-z}$.

2 Consider the leading order coefficient of x in the numerator of the expansion of $\prod_{i=1}^{N} \frac{x^k}{y^2 + (x-x_i)^2}$. In particular, $\sum_{i=1}^{N} C_{i+}(x_{i+})^k + C_{i-}(x_{i-})^k = 0$ for $k = 0, 1, \ldots, 2(N-1)$.

3 Write $\mathrm{Log}(r)^{\pm} = \mathrm{Log}\, r(1+\varepsilon^{\pm})$ and note that $\varepsilon^{\pm} \to 0$ for $b \to \infty$, $a \to -\infty$ respectively.

4 The binomial series for non-integer α is given by $(1+x)^{\alpha} = \sum_{i=0}^{\infty} \frac{\Gamma(\alpha+1)}{\Gamma(\alpha-i+1)\, i!} x^i$, where $|x| < 1$.

5 The Exponential Integral is given by $\mathrm{Ei}(z) = \int_{-\infty}^{z} \frac{e^u}{u} du$ where the path of integration does not cut $[0, \infty)$. It may be written as $\mathrm{Ei}(z) = \gamma + \mathrm{Log}(-z) + \sum_{i=1}^{\infty} \frac{z^i}{i\, i!}$ in the plane cut along the positive real axis.

APPENDIX

The following is a Mathematica program for finding the kth posterior moment for the N observations $x_1,,().\ x_N$ given position $y_0 = y$ and uniform prior on x $\infty{-}\infty,($.

```
kthMoment[k_Integer,N_Integer] := kthIntegral[k,N] / kthIntegral[0,N],
```

where

```
kthIntegral[k_Integer,N_Integer] :=
Block[      {i,j,Result},
            For[i=1,i<=N,i++,
                  xp[i] = x[i] + I y;
                  xm[i] = x[i] - I y;
            ];

            For[i=1,i<=N,i++,
                  Cp[i] =  1/(xp[i]-xm[i]);
                  Cm[i] = -1/(xp[i]-xm[i]);
                  For[j=1,j<=N,j++,
                        If[ j!=i,
                              Cp[i] *= 1/(xp[i]-xp[j]) 1/(xp[i]-xm[j]);
                              Cm[i] *= 1/(xm[i]-xp[j]) 1/(xm[i]-xm[j]);
                        ];
                  ];
            ];

            Result = 0;

            For[i=1,i<=N,i++,
                  Result += Cp[i] xp[i]^k - Cm[i] xm[i]^k;
            ];

            Return[Result];
]
```