Hypothesis Testing for Shape Parameter of Gamma Distribution

Anil Shanbhag   Sarthak Bagaria

Hypothesis

$$H_0 : k = 5$$

vs.

$$H_1 : k > 5$$

where k is the shape parameter of the Gamma distribution from which a sample of 50 points is taken.

Test Statistic

- For gamma distribution (GM, AM) is a minimal sufficient statistic for (k,β).
- Based on the paper by Keating et al.[1] we use the test statistic w:

$$
\begin{aligned}
GM &= (\prod_{i=1}^{n} T_i)^{1/n} \\
AM &= \sum_{i=1}^{n} T_i/n \\
w &= GM/AM
\end{aligned}
$$

  where T_i are the values from a gamma sample.
- (w, AM) is also minimally sufficient for (k, β) and the pair is stochastically independent.

Gamma (5,1) Sample

2.6283619209343	7.69450613960945	6.93695507619144	3.40969289304186
5.01595406280265	2.17678469169354	8.80340669389395	6.92950635515724
3.28724585442206	1.09957908734132	6.2034202591892	7.27085416696607
7.8356080618422	6.41677034478504	7.75632623464187	4.9645397847155
4.38203553112516	5.74649177978731	3.8166862081331	4.83586834774907
4.21534304747825	2.76046261953365	2.6833092226359	6.14388003733696
2.76639643059915	5.95142981282601	10.9850460828745	7.03720119143969
4.44532021982175	6.60770695474903	2.55801016908204	2.44764702391154
5.74621735480312	2.7025937057668	2.5400119096503	5.92742255244166
3.50569807423671	4.26590769266725	7.41203931734455	4.65915398021402
2.83258965327326	3.08623349517042	7.38523921929653	6.51932172633605
2.59312896786055	5.87443896213083	3.78131726237567	5.54589735520884
4.72629034438013	6.3652634420665		

- For our sample, we obtain the value w = 0.909968

Critical Region

- This is the region within which if obtained value of w lies, we reject the null-hypothesis.
- $w_c = w_{(\alpha,k,n)}$ is the critical value for α confidence level.
- Critical region $CR = \{w : w > w_c\}$
- For $n > 30$ we approximate the distribution to Beta distribution by employing technique by Patnaik (1949).
- In this technique we equate the first two moments of W and Beta distribution to find the parameters α and β of Beta distribution.

Critical Region (Contd.)

- Equations to solve :

$$\alpha/(\alpha+\beta)$$
$$= [\Gamma(k+1/n)/\Gamma(k)]^{n-1} \times \prod_{i=1}^{n-1} [\Gamma(k+i/n)/\Gamma(k+1/n+i/n)]$$

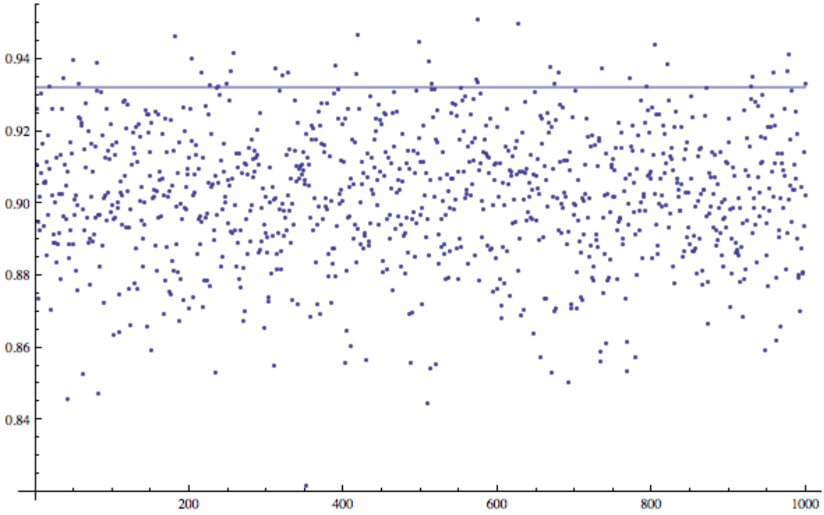
$$\alpha(\alpha+1)/[(\alpha+\beta)(\alpha+\beta+1)]$$
$$= [\Gamma(k+2/n)/\Gamma(k)]^{n-1} \times \prod_{i=1}^{n-1} [\Gamma(k+i/n)/\Gamma(k+2/n+i/n)]$$

- For $k = 5$ and $n = 50$ with level of significance 0.05 we get $w_c = 0.932226$

Hypothesis Testing

- We see that $w < w_c$. Hence for our sample we accept the null hypothesis.
- Also $w_{(0.5,5,50)} = 0.904888$, the median for null distribution is quite close to our observed w.
- For 1000 simulations, we found our hypothesis to be rejected only 55 times i.e. 0.055 in fraction.

1000 Samples Experiment



Median Unbiased Estimator

- The MU estimator for k^* is such that $w = w_{(0.5,k^*,n)}$.
- By solving the above equations taking α, β, k as variables, we obtain $k^* = 5.28736$, which we see to be quite close to value used for sample generation.

Confidence Bound

- Solving for k' in $w_{(0.05,k',n)} = w$, using same method as above, we obtain the 95% confidence bound.
- For all κ above this value, the null hypothesis $H_0 : k = \kappa$ is not rejected for the alternative $H_1 : k > \kappa$
- For our sample $\kappa = 3.7599$

References

Keating J.P., Glaser R.E, Ketchum N.S. *Testing Hypothesis About the Shape Parameter of a Gamma Distribution*. Technometrics, February 1990, Vol. 32, No. 1

```
wl = {}
k = 5
n = 50
sample = RandomVariate[GammaDistribution[5, 1], n];
w = GeometricMean[sample] / Mean[sample];
sols =
 N[
  Solve[
   {a / (a + b) ==
     ((Gamma[k + 1 / n] / Gamma[k]) ^ (n - 1) *
       Product[Gamma[k + i / n] / Gamma[k + 1 / n + i / n], {i, 1, n - 1}]),
    a * (a + 1) / ((a + b) * (a + b + 1)) ==
     ((Gamma[k + 2 / n] / Gamma[k]) ^ (n - 1) *
       Product[Gamma[k + i / n] / Gamma[k + 2 / n + i / n], {i, 1, n - 1}])},
   {a, b}]
 ]

InverseCDF[BetaDistribution[a /. sols[[1]], b /. sols[[1]]],
 0.95]
```

```
wl = {}
k = 5
n = 50
Do[
    sample = RandomVariate[GammaDistribution[5, 1], n];
    w = GeometricMean[sample] / Mean[sample];
  AppendTo[wl, w]
 , {i, 1000}]
wl
sample = RandomVariate[GammaDistribution[5, 1], n];
w = GeometricMean[sample] / Mean[sample];
sols =
 N[
  Solve[
   {a / (a + b) ==
      ((Gamma[k + 1 / n] / Gamma[k]) ^ (n - 1) *
        Product[Gamma[k + i / n] / Gamma[k + 1 / n + i / n], {i, 1, n - 1}]),
    a * (a + 1) / ((a + b) * (a + b + 1)) ==
      ((Gamma[k + 2 / n] / Gamma[k]) ^ (n - 1) *
        Product[Gamma[k + i / n] / Gamma[k + 2 / n + i / n], {i, 1, n - 1}])}, {a, b}]
 ]

InverseCDF[BetaDistribution[a /. sols[[1]], b /. sols[[1]]], 0.5]
cutoff = InverseCDF[BetaDistribution[a /. sols[[1]], b /. sols[[1]]], 0.95]
gc1 = ListPlot[wl]
gc2 = Plot[cutoff, {x, 0, 1000}]
Show[gc1, gc2]
```