Cross Validated

Entropy of the multinomial distribution

Asked 5 years, 3 months ago Active 2 years, 4 months ago Viewed 4k times

7

In my work I've found myself in the position of needing to calculate the entropy of the multinomial distribution:

$$\text{Multinomial}(\mathbf{x};\ \ n, \mathbf{p})$$

I imagine it would be too much to expect a closed-form formula for this value, but what is the current standard method for efficiently calculating/approximating the entropy of a multinomial distribution?

Also my situation is for small n, so asymptotic approximations probably aren't the best for me unless they converge extremely quickly.

multinomial-distribution entropy

Share Cite Improve this question Follow

edited Apr 18 '16 at 20:41

asked Apr 18 '16 at 4:54

Thoth
1,195 7 24

1 Answer

Active Oldest Votes

6

Ok so I guess I should have done a bit of experimentation before posting this question. I just assumed that since the Wikipedia article for the multinomial distribution didn't mention entropy, and since I couldn't find anything about it on google, that it was very difficult to compute.

Let

$$\mathbf{X} \sim \text{Multinomial}(n, \mathbf{p})$$

The entropy for $\mathbf{X}$ is given by:

$$\mathrm{H}(\mathbf{X}) = -\sum_{\substack{\mathbf{x}\geq 0 \\ \sum_i x_i = n}} \frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} \log \left[\frac{n!}{x_1! \cdots x_k!} p_1^{x_1} \cdots p_k^{x_k} \right].$$

Using the logarithm to break this up we obtain:

$$\begin{aligned}
\mathrm{H}(\mathbf{X}) &= -\log n! - \sum_{i=1}^{k} \log p_i \mathrm{E}[X_i] + \sum_{i=1}^{k} \mathrm{E}[\log X_i!] \\
&= -\log n! - n \sum_{i=1}^{k} p_i \log p_i + \sum_{i=1}^{k} \mathrm{E}[\log X_i!] \\
&= -\log n! - n \sum_{i=1}^{k} p_i \log p_i + \sum_{i=1}^{k} \sum_{x_i=0}^{n} \binom{n}{x_i} p_i^{x_i} (1-p_i)^{n-x_i} \log x_i!.
\end{aligned}$$

Thus we see that instead of summing over all distinct permutations of the partitions of n, which scales exponentially with both the size of n and k, the derived form scales as $O((n+1)k)$, which is linear in both n and k.

Share Cite Improve this answer Follow

edited Mar 31 '19 at 23:28

answered Apr 19 '16 at 0:21

Thoth
1,195 7 24

Two Q's: 1) Where did $E[X_i]$ come from? 2) From eq. (1) to (2), where did everything before log go? – user13985 Sep 16 '19 at 0:49

1 @user13985 recall the definition of entropy is just $\mathrm{E}[-\log(P(X))]$, so when the log breaks things up you just get a bunch of simpler expected values. The stuff before the log is just the probability distribution itself (multinomial pmf) and so it is just folded into the definition of expected value. – Thoth Sep 16 '19 at 1:18

I see! I was trying to use this definition $H(X) = -\sum p(x) \log p(x)$. Instead, I should use this definition $H(X) = E[-\log(P(X))]$. I think that's the problem? – user13985 Sep 16 '19 at 1:31

@user13985, well the former definition is canon. I provided you with the alternative definition to help make the point that it can be interpreted as an expected value, and to emphasize that it's a common pattern to look out for in the future, since once you recognize that an expression can be interpreted as an expected value, it may allow you to better see how to solve the problem. – Thoth Sep 16 '19 at 1:41

1 @user13985 I'm not sure I understand your concern, the definitions are equal to each other, you can switch back and forth between them at will. If you break things up with the log prior to interpreting it as an expected value, you can then at the next step just interpret each of the pieces as an expected value, they all have the appropriate form, i.e. $\sum_x p(x)f(x)$. – Thoth Sep 16 '19 at 1:46