**Gartner-Ellis Theorem**

Consider for example N random, Gaussian distributed variables, Xi, with mean μ, and standard deviation σ. Then the variable:



is also Gaussian distributed with mean μ and standard deviation σ/√n. In other words, its pdf is:



We can write the dominant part of the distribution as:



where ψ(x) is called the rate function. Consider also the sum of Poisson distributed variables, Xi, with mean μ (the std is related to μ I think I remember). Anyway, now let’s consider the distribution of N, defined as:



It is approximately given by:



So it appears that, ignoring the fluctuations, we can write the pdf of a random variable, N, as:



So the question falls to that of determining what the rate function is. We can sort of derive it through the following considerations. Consider a general N-variable probability distribution PN(**x**), and suppose we wish to calculate the probability distribution of some function of the random variables, sN(**X**). Then we wish to calculate:



We will use the Laplace representation of the δ function; it is just the inverse Laplace transform of 1, which is:



Then we can write:



We can calculate the expectation integral on the right and obtain, perhaps, something like,



where the last term is something that goes to 0 as N goes to ∞. Substituting this into the integral we have:



where in the last we readjust what a is to get rid of the Np divisor and also ignore the prefactor Np since we’re just looking for dominant contributions. Now we’re in position to do a saddle point approximation to the integral. We deform the contour so that it goes through the saddle point of ζs – λ(ζ), along the path of steepest descent. Let ζ \* be the saddle point. Then we have:



So therefore, in the limit that N → ∞, we have:



which is the Gartner-Ellis theorem.

**Example**

Let’s consider the Gaussian random variables again for example. Then,



Continuing then, we have:



and so then using the saddle point approximation, though of course this can be done exactly, we have:



**Example**

Let’s consider the CLT. I’ll have to take Fourrier transform rather, since the Laplace won’t exist for all distributions. And consider the variable they use: S/√N.



At this point we would endeavor an approximation via Taylor series:



And then we could say:



If we change variables to s, then we’d have:



Which is correct. But note that if we tried to calculate p(s) directly this way, we would’ve run into a problem,



Since the Taylor series wouldn’t obviously converge to anything:



And so there is no reason to stop. And consider if we had done a little differently, and considered S/N to begin with:



And then,



But it seems that we might just get rid of the variance term entirely, since it would disappear in the limit. Let’s reprise the example, using the saddle point approximation:



Let’s consider a specific probability distribution. Let’s take a uniform distribution between [-1,1]. Then,



And so then we evaluate:



Attempting an SPA is problematic since sin(ζ) goes to zero and negative. The exponential, even if N were even, would be riddled with poles. But I could close the integral in the complex plane into the u.h.p., as long as s were greater than N? But s can never be greater than N. So saddle point approximation is probably not going to work.