Probability Theory

1. *Characteristic Functions:*

The characteristic function is defined by

$$f(k) \equiv \langle \exp(-ikx) \rangle = \int \exp(-ikx) p(x) dx.$$

The *nth* coefficient of the Taylor series of $f(k)$, expanded around $k = 0$, gives the *nth* moment of x as

$$f(k) = \sum_{n=0}^{\infty} \frac{(-ik)^n}{n!} \langle x^n \rangle.$$

(a) A *uniform* probability distribution,

$$p(x) = \begin{cases} \dfrac{1}{2a} & \text{for } -a < x < a \\ 0 & \text{otherwise} \end{cases},$$

for which there exist many examples, gives

$$f(k) = \frac{1}{2a} \int_{-a}^{a} \exp(-ikx) dx = \frac{1}{2a} \frac{1}{-ik} \exp(-ikx) \Big|_{-a}^{a}$$

$$= \frac{1}{ak} \sin(ka) = \sum_{m=0}^{\infty} (-1)^m \frac{(ak)^{2m}}{(2m+1)!}.$$

Therefore,

$$m_1 = \langle x \rangle = 0, \qquad \text{and} \qquad m_2 = \langle x^2 \rangle = \frac{1}{3} a^2.$$

(b) The *Laplace* PDF,

$$p(x) = \frac{1}{2a} \exp\left(-\frac{|x|}{a}\right),$$

for example describing light absorption through a turbid medium, gives

$$\begin{aligned} f(k) &= \frac{1}{2a} \int_{-\infty}^{\infty} dx \exp\left(-ikx - \frac{|x|}{a}\right) \\ &= \frac{1}{2a} \int_{0}^{\infty} dx \exp(-ikx - x/a) + \frac{1}{2a} \int_{-\infty}^{0} dx \exp(-ikx + x/a) \\ &= \frac{1}{2a} \left[\frac{1}{-ik + 1/a} - \frac{1}{-ik - 1/a}\right] = \frac{1}{1 + (ak)^2} \\ &= 1 - (ak)^2 + (ak)^4 - \cdots. \end{aligned}$$

Therefore,

$$m_1 = \langle x \rangle = 0, \qquad \text{and} \qquad m_2 = \langle x^2 \rangle = 2a^2.$$

(c) The *Cauchy*, or *Lorentz* PDF describes the spectrum of light scattered by diffusive modes, and is given by

$$p(x) = \frac{a}{\pi(x^2 + a^2)}.$$

For this distribution,

$$\begin{aligned} f(k) &= \int_{-\infty}^{\infty} \exp(-ikx) \frac{a}{\pi(x^2 + a^2)} dx \\ &= \frac{1}{2\pi i} \int_{-\infty}^{\infty} \exp(-ikx) \left[\frac{1}{x - ia} - \frac{1}{x + ia} \right] dx. \end{aligned}$$

The easiest method for evaluating the above integrals is to close the integration contours in the complex plane, and evaluate the residue. The vanishing of the integrand at infinity determines whether the contour has to be closed in the upper, or lower half of the complex plane, and leads to

$$f(k) = \left\{ \begin{array}{ll} -\dfrac{1}{2\pi i} \displaystyle\int_C \frac{\exp(-ikx)}{x + ia} dx = \exp(-ka) & \text{for } k \geq 0 \\ \dfrac{1}{2\pi i} \displaystyle\int_B \frac{\exp(-ikx)}{x - ia} dx = \exp(ka) & \text{for } k < 0 \end{array} \right\} = \exp(-|ka|).$$

Note that $f(k)$ is not an analytic function in this case, and hence does not have a Taylor expansion. The moments have to be determined by another method, e.g. by direct evaluation, as

$$m_1 = \langle x \rangle = 0, \qquad \text{and} \qquad m_2 = \langle x^2 \rangle = \int dx \frac{\pi}{a} \cdot \frac{x^2}{x^2 + a^2} \to \infty.$$

The first moment vanishes by symmetry, while the second (and higher) moments diverge, explaining the non-analytic nature of $f(k)$.

(d) The *Rayleigh* distribution,

$$p(x) = \frac{x}{a^2} \exp\left(-\frac{x^2}{2a^2} \right), \qquad \text{for} \qquad x \geq 0,$$

can be used for the length of a random walk in two dimensions. Its characteristic function is

$$\begin{aligned} f(k) &= \int_0^{\infty} \exp(-ikx) \frac{x}{a^2} \exp\left(-\frac{x^2}{2a^2} \right) dx \\ &= \int_0^{\infty} [\cos(kx) - i \sin(kx)] \frac{x}{a^2} \exp\left(-\frac{x^2}{2a^2} \right) dx. \end{aligned}$$

The integrals are not simple, but can be evaluated as

$$\int_0^\infty \cos(kx)\frac{x}{a^2}\exp\left(-\frac{x^2}{2a^2}\right)dx = \sum_{n=0}^{\infty}\frac{(-1)^n n!}{(2n)!}\left(2a^2k^2\right)^n,$$

and

$$\begin{aligned}\int_0^\infty \sin(kx)\frac{x}{a^2}\exp\left(-\frac{x^2}{2a^2}\right)dx &= \frac{1}{2}\int_{-\infty}^\infty \sin(kx)\frac{x}{a^2}\exp\left(-\frac{x^2}{2a^2}\right)dx\\ &= \sqrt{\frac{\pi}{2}}\,ka\,\exp\left(-\frac{k^2a^2}{2}\right),\end{aligned}$$

resulting in

$$f(k) = \sum_{n=0}^{\infty}\frac{(-1)^n n!}{(2n)!}\left(2a^2k^2\right)^n - i\sqrt{\frac{\pi}{2}}\,ka\,\exp\left(-\frac{k^2a^2}{2}\right).$$

The moments can also be calculated directly, from

$$\begin{aligned}m_1 = \langle x\rangle &= \int_0^\infty \frac{x^2}{a^2}\exp\left(-\frac{x^2}{2a^2}\right)dx = \int_{-\infty}^\infty \frac{x^2}{2a^2}\exp\left(-\frac{x^2}{2a^2}\right)dx = \sqrt{\frac{\pi}{2}}a,\\ m_2 = \langle x^2\rangle &= \int_0^\infty \frac{x^3}{a^2}\exp\left(-\frac{x^2}{2a^2}\right)dx = 2a^2\int_0^\infty \frac{x^2}{2a^2}\exp\left(-\frac{x^2}{2a^2}\right)d\left(\frac{x^2}{2a^2}\right)\\ &= 2a^2\int_0^\infty y\exp(-y)dy = 2a^2.\end{aligned}$$

(e) It is difficult to calculate the characteristic function for the *Maxwell* distribution

$$p(x) = \sqrt{\frac{2}{\pi}}\frac{x^2}{a^3}\exp\left(-\frac{x^2}{2a^2}\right),$$

say describing the speed of a gas particle. However, we can directly evaluate the mean and variance, as

$$\begin{aligned}m_1 = \langle x\rangle &= \sqrt{\frac{2}{\pi}}\int_0^\infty \frac{x^3}{a^3}\exp\left(-\frac{x^2}{2a^2}\right)dx\\ &= 2\sqrt{\frac{2}{\pi}}a\int_0^\infty \frac{x^2}{2a^2}\exp\left(-\frac{x^2}{2a^2}\right)d\left(\frac{x^2}{2a^2}\right)\\ &= 2\sqrt{\frac{2}{\pi}}a\int_0^\infty y\exp(-y)dy = 2\sqrt{\frac{2}{\pi}}a,\end{aligned}$$

and

$$m_2 = \langle x^2\rangle = \sqrt{\frac{2}{\pi}}\int_o^\infty \frac{x^4}{a^3}\exp\left(-\frac{x^2}{2a^2}\right)dx = 3a^2.$$

****

2. At each step a *directed random walk* can move along angles θ and ϕ with probability

$$p(\theta) = \frac{2}{\pi}\cos^2\left(\frac{\theta}{2}\right), \qquad \text{and} \qquad p(\phi) = \frac{1}{2\pi},$$

where the solid angle factor of $\sin\theta$ is already included in the definition of $p(\theta)$;

$$\int p(\theta)d\theta = \int_0^\pi \frac{2}{\pi}\cos^2\left(\frac{\theta}{2}\right)d\theta = \int_0^\pi \frac{\cos\theta + 1}{\pi}d\theta = 1.$$

(a) From symmetry arguments,

$$\langle x \rangle = \langle y \rangle = 0,$$

while along the z-axis,

$$\langle z \rangle = \sum_i \langle z_i \rangle = N\langle z_i \rangle = Na\left\langle \cos\theta_i \right\rangle = \frac{Na}{2}.$$

The last equality follows from

$$\begin{aligned}\left\langle \cos\theta_i \right\rangle &= \int p(\theta)\cos\theta d\theta = \int_0^\pi \frac{1}{\pi}\cos\theta\cdot(\cos\theta+1)d\theta \\ &= \int_0^\pi \frac{1}{2\pi}(\cos 2\theta + 1)d\theta = \frac{1}{2}.\end{aligned}$$

The second moment of z is given by

$$\begin{aligned}\left\langle z^2 \right\rangle &= \sum_{i,j}\left\langle z_i z_j \right\rangle = \sum_i \sum_{i\neq j}\left\langle z_i z_j \right\rangle + \sum_i \left\langle z_i^2 \right\rangle \\ &= \sum_i \sum_{i\neq j}\left\langle z_i \right\rangle \left\langle z_j \right\rangle + \sum_i \left\langle z_i^2 \right\rangle \\ &= N(N-1)\left\langle z_i \right\rangle^2 + N\left\langle z_i^2 \right\rangle.\end{aligned}$$

Noting that

$$\frac{\left\langle z_i^2 \right\rangle}{a^2} = \int_0^\pi \frac{1}{\pi}\cos^2\theta(\cos\theta+1)d\theta = \int_0^\pi \frac{1}{2\pi}(\cos 2\theta + 1)d\theta = \frac{1}{2},$$

we find

$$\left\langle z^2 \right\rangle = N(N-1)\left(\frac{a}{2}\right)^2 + N\frac{a^2}{2} = N(N+1)\frac{a^2}{4}.$$

The second moments in the x and y directions are equal, and given by

$$\left\langle x^2 \right\rangle = \sum_{i,j}\left\langle x_i x_j \right\rangle = \sum_i \sum_{i\neq j}\left\langle x_i x_j \right\rangle + \sum_i \left\langle x_i^2 \right\rangle = N\left\langle x_i^2 \right\rangle.$$

Using the result

$$\begin{aligned}\frac{\langle x_i^2\rangle}{a^2} &= \langle \sin^2\theta\cos^2\phi\rangle \\ &= \frac{1}{2\pi^2}\int_0^{2\pi} d\phi\cos^2\phi\int_0^{\pi} d\theta\sin^2\theta(\cos\theta+1) = \frac{1}{4},\end{aligned}$$

we obtain

$$\langle x^2\rangle = \langle y^2\rangle = \frac{Na^2}{4}.$$

While the variables x, y, and z are not independent because of the constraint of unit length, simple symmetry considerations suffice to show that the three covariances are in fact zero, i.e.

$$\langle xy\rangle = \langle xz\rangle = \langle yz\rangle = 0.$$

(b) From the Central limit theorem, the probability density should be Gaussian. However, for correlated random variable we may expect cross terms that describe their covariance. Since we showed above that the covaraince between x, y, and z are all zero, we can treat them as three *independent* Gaussian variables, and write

$$p(x,y,z)\propto\exp\left[-\frac{(x-\langle x\rangle)^2}{2\sigma_x^2}-\frac{(y-\langle y\rangle)^2}{2\sigma_y^2}-\frac{(z-\langle z\rangle)^2}{2\sigma_z^2}\right].$$

(There will be correlations between x, y, and z appearing in higher cumulants, but all such cumulants become irrelevant in the $N\to\infty$ limit.) Using the moments

$$\langle x\rangle = \langle y\rangle = 0, \qquad \text{and} \qquad \langle z\rangle = N\frac{a}{2},$$

$$\sigma_x^2 = \langle x^2\rangle - \langle x\rangle^2 = N\frac{a^2}{4} = \sigma_y^2,$$

$$\text{and} \qquad \sigma_z^2 = \langle z^2\rangle - \langle z\rangle^2 = N(N+1)\frac{a^2}{4} - \left(\frac{Na}{2}\right)^2 = N\frac{a^2}{4},$$

we obtain

$$p(x,y,z) = \left(\frac{2}{\pi Na^2}\right)^{3/2}\exp\left[-\frac{x^2+y^2+(z-Na/2)^2}{Na^2/2}\right].$$

****

3. *Tchebycheff's Inequality:* By definition, for a system with a PDF $p(x)$, and average λ, the variance is

$$\sigma^2 = \int (x-\lambda)^2 p(x)dx.$$

Let us break the integral into two parts as

$$\sigma^2 = \int_{|x-\lambda| \geq n\sigma} (x-\lambda)^2 p(x) dx + \int_{|x-\lambda| < n\sigma} (x-\lambda)^2 p(x) dx,$$

resulting in

$$\sigma^2 - \int_{|x-\lambda| < n\sigma} (x-\lambda)^2 p(x) dx = \int_{|x-\lambda| \geq n\sigma} (x-\lambda)^2 p(x) dx.$$

Now since

$$\int_{|x-\lambda| \geq n\sigma} (x-\lambda)^2 p(x) dx \geq \int_{|x-\lambda| \geq n\sigma} (n\sigma)^2 p(x) dx,$$

we obtain

$$\int_{|x-\lambda| \geq n\sigma} (n\sigma)^2 p(x) dx \leq \sigma^2 - \int_{|x-\lambda| < n\sigma} (x-\lambda)^2 p(x) dx \leq \sigma^2,$$

and

$$\int_{|x-\lambda| \geq n\sigma} p(x) dx \leq \frac{1}{n^2}.$$

****

4. *Optimal Selections*:

(a) The probability that the maximum of n random numbers falls between x and $x + dx$ is equal to the probability that one outcome is in this interval, while all the others are smaller than x, i.e.

$$p_n(x) = p(r_1 = x, r_2 < x, r_3 < x, \cdots, r_n < x) \times \begin{pmatrix} n \\ 1 \end{pmatrix},$$

where the second factor corresponds to the number of ways of choosing which $r_\alpha = x$. As these events are independent

$$\begin{aligned} p_n(x) &= p(r_1 = x) \cdot p(r_2 < x) \cdot p(r_3 < x) \cdots p(r_n < x) \times \binom{n}{1} \\ &= p(r = x) \left[p(r < x)\right]^{n-1} \times \binom{n}{1}. \end{aligned}$$

The probability of $r < x$ is just a cumulative probability function, and

$$p_n(x) = n \cdot p(x) \cdot \left[\int_0^x p(r) dr\right]^{n-1}.$$

(b) If each r_α is uniformly distributed between 0 and 1, $p(r) = 1$ ($\int_0^1 p(r)dr = \int_0^1 dr = 1$). With this PDF, we find

$$p_n(x) = n \cdot p(x) \cdot \left[\int_0^x p(r)dr\right]^{n-1} = n\left[\int_0^x dr\right]^{n-1} = nx^{n-1},$$

and the mean is now given by

$$\langle x \rangle = \int_0^1 x p_n(x)dx = n\int_0^1 x^n dx = \frac{n}{n+1}.$$

The second moment of the maximum is

$$\left\langle x^2 \right\rangle = n\int_0^1 x^{n+1}dx = \frac{n}{n+2},$$

resulting in a variance

$$\sigma^2 = \left\langle x^2 \right\rangle - \langle x \rangle^2 = \frac{n}{n+2} - \left(\frac{n}{n+1}\right)^2 = \frac{n}{(n+1)^2(n+2)}.$$

Note that for large n the mean approaches the limiting value of unity, while the variance vanishes as $1/n^2$. There is too little space at the top of the distribution for a wide variance.

****

5. *Information:*

(a) For an unbiased probability estimation, we need to maximize entropy subject to the two constraints of normalization, and of given average speed ($\langle |v| \rangle = c.$). Using Lagrange multipliers α and β to impose these constraints, we need to maximize

$$S = -\langle \ln p \rangle = -\int_{-\infty}^{\infty} p(v)\ln p(v)dv + \alpha\left(1 - \int_{-\infty}^{\infty} p(v)dv\right) + \beta\left(c - \int_{-\infty}^{\infty} p(v)|v|dv\right).$$

Extremizing the above expression yields

$$\frac{\partial S}{\partial p(v)} = -\ln p(v) - 1 - \alpha - \beta|v| = 0,$$

which is solved for

$$\ln p(v) = -1 - \alpha - \beta|v|,$$

or

$$p(v) = Ce^{-\beta|v|}, \qquad \text{with} \quad C = e^{-1-\alpha}.$$

The constraints can now be used to fix the parameters C and β:

$$1 = \int_{-\infty}^{\infty} p(v)dv = \int_{-\infty}^{\infty} Ce^{-\beta|v|}dv = 2C \int_{0}^{\infty} e^{-\beta v}dv = 2C\frac{1}{-\beta}e^{-\beta v}\Big|_{0}^{\infty} = \frac{2C}{\beta},$$

which implies

$$C = \frac{\beta}{2}.$$

From the second constraint, we have

$$c = \int_{-\infty}^{\infty} Ce^{-\beta|v|}|v|dv = \beta \int_{0}^{\infty} e^{-\beta v}vdv,$$

which when integrated by parts yields

$$c = \beta\left[-\frac{1}{\beta}ve^{-\beta v}\Big|_{0}^{\infty} + \frac{1}{\beta}\int_{0}^{\infty} e^{-\beta v}dv\right] = \left[-\frac{1}{\beta}\, e^{-\beta v}\big|_{0}^{\infty}\right] = \frac{1}{\beta}.$$

or,

$$\beta = \frac{1}{c}.$$

The unbiased PDF is then given by

$$p(v) = Ce^{-\beta|v|} = \frac{1}{2c}\exp\left(-\frac{|v|}{c}\right).$$

(b) When the second constraint is on the average kinetic energy, $\langle mv^2/2\rangle = mc^2/2$, we have

$$S = -\int_{-\infty}^{\infty} p(v)\ln p(v)dv + \alpha\left(1 - \int_{-\infty}^{\infty} p(v)dv\right) + \beta\left(\frac{mc^2}{2} - \int_{-\infty}^{\infty} p(v)\frac{mv^2}{2}dv\right).$$

The corresponding extremization,

$$\frac{\partial S}{\partial p(v)} = -\ln p(v) - 1 - \alpha - \beta\frac{mv^2}{2} = 0,$$

results in

$$p(v) = C\exp\left(-\frac{\beta mv^2}{2}\right).$$

The normalization constraint implies

$$1 = \int_{-\infty}^{\infty} p(v)dv = C\int_{-\infty}^{\infty} e^{-\beta mv^2/2} = C\sqrt{2\pi/\beta m},$$

or

$$C = \sqrt{\beta m / 2\pi}.$$

The second constraint,

$$\begin{aligned}
\frac{mc^2}{2} &= \int_{-\infty}^{\infty} p(v) \frac{mv^2}{2} dv = \frac{m}{2}\sqrt{\frac{\beta m}{2\pi}} \int_{-\infty}^{\infty} \exp\left(-\frac{\beta m v^2}{2}\right) v^2 dv \\
&= \frac{m}{2}\sqrt{\frac{\beta m}{2\pi}} \left[\frac{\sqrt{\pi}}{2}\left(\frac{2}{\beta m}\right)^{3/2}\right] = \frac{1}{2\beta},
\end{aligned}$$

gives

$$\beta = \frac{1}{mc^2},$$

for a full PDF of

$$p(v) = C \exp\left(-\frac{\beta m v^2}{2}\right) = \frac{1}{\sqrt{2\pi c^2}} \exp\left(-\frac{v^2}{2c^2}\right).$$

(c) The entropy of the first PDF is given by

$$\begin{aligned}
S_1 = -\langle \ln p_1 \rangle &= -\int_{-\infty}^{\infty} \frac{1}{2c} \exp\left(\frac{-|v|}{c}\right) \left[-\ln(2c) - \frac{|v|}{c}\right] dv \\
&= \frac{\ln(2c)}{c} \int_0^{\infty} \exp\left(-\frac{v}{c}\right) dv + \frac{1}{c}\int_0^{\infty} \exp\left(-\frac{v}{c}\right) \frac{v}{c} dv \\
&= -\ln(2c) \exp(-v/c)\big|_0^{\infty} + \frac{1}{c}\left[-c \exp(-v/c)\big|_0^{\infty}\right] \\
&= \ln(2c) + 1 = 1 + \ln 2 + \ln c.
\end{aligned}$$

For the second distribution, we obtain

$$\begin{aligned}
S_2 = -\langle \ln p_2 \rangle &= -\frac{1}{\sqrt{2\pi c^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{v^2}{2c^2}\right) \left[-\frac{1}{2}\ln\left(2\pi c^2\right) - \frac{v^2}{2c^2}\right] \\
&= \frac{\ln\left(2\pi c^2\right)}{2\sqrt{2\pi c^2}} \int_{-\infty}^{\infty} \exp\left(-v^2/2c^2\right) dv + \frac{1}{\sqrt{2\pi c^2}} \int_{-\infty}^{\infty} \frac{v^2}{2c^2} \exp\left(-v^2/2c^2\right) dv \\
&= \frac{1}{2}\ln\left(2\pi c^2\right) + \frac{1}{c^2\sqrt{2\pi c^2}} \left[\frac{\sqrt{2\pi c^2}\, c^2}{2}\right] \\
&= \frac{1}{2}\ln\left(2\pi c^2\right) + \frac{1}{2} = \frac{1}{2} + \frac{1}{2}\ln(2\pi) + \ln c.
\end{aligned}$$

For a discrete probability, the information content is

$$I_\alpha = \ln_2 M - S_\alpha / \ln 2,$$

where M denotes the number of possible outcomes. While M, and also the proper measure of probability are not well defined for a continuous PDF, the ambiguities disappear when we consider the difference

$$\begin{aligned} I_2 - I_1 &= \left(-S_2 + S_1\right) / \ln 2 \\ &= -\left(S_2 - S_1\right) / \ln 2 \\ &= -\frac{(\ln \pi - \ln 2 - 1)}{2 \ln 2} \approx 0.3956. \end{aligned}$$

Hence the constraint of constant energy provides 0.3956 more bits of information. (This is partly due to the larger variance of the distribution with constant speed.)

****

6. *Benford's Law:* Let us consider the observation that the probability distribution for first integers is unchanged under multiplication by any (i.e. a random) number. Presumably we can repeat such multiplications many times, and it is thus suggestive that we should consider the properties of the *product of random numbers.* (Why this should be a good model for stock prices is not entirely clear, but it seems to be as good an explanation as anything else!)

Consider the $x = \prod_{i=1}^{N} r_i$, where r_i are *positive*, random variables taken from some reasonably well behaved probability distribution. The random variable $\ell \equiv \ln x = \sum_{i=1}^{N} \ln r_i$ is the sum of many random contributions, and according to the central limit theorem should have a Gaussian distribution in the limit of large N, i.e.

$$\lim_{N \to \infty} p(\ell) = \exp\left[-\frac{\left(\ell - N\overline{\ell}\right)^2}{2N\sigma^2}\right] \frac{1}{\sqrt{2\pi N\sigma^2}},$$

where $\overline{\ell}$ and σ^2 are the mean and variance of $\ln r$ respectively. The product x is distributed according to the *log-normal distribution*

$$p(x) = p(\ell)\frac{d\ell}{dx} = \frac{1}{x}\exp\left[-\frac{\left(\ln(x) - N\overline{\ell}\right)^2}{2N\sigma^2}\right] \frac{1}{\sqrt{2\pi N\sigma^2}}.$$

The probability that the first integer of x in a decimal representation is i is now obtained approximately as follows:

$$p_i = \sum_q \int_{10^q i}^{10^q (i+1)} dx p(x).$$

The integral considers cases in which x is a number of magnitude 10^q (i.e. has $q+1$ digits before the decimal point). Since the number is quite widely distributed, we then have to sum over possible magnitudes q. The range of the sum actually need not be specified! The next stage is to change variables from x to $\ell = \ln x$, leading to

$$p_i = \sum_q \int_{q+\ln i}^{q+\ln(i+1)} d\ell p(\ell) = \sum_q \int_{q+\ln i}^{q+\ln(i+1)} d\ell \exp\left[-\frac{\left(\ell - N\overline{\ell}\right)^2}{2N\sigma^2}\right] \frac{1}{\sqrt{2\pi N\sigma^2}}.$$

We shall now make the *approximation* that over the range of integration ($q + \ln i$ to $q + \ln(i+1)$), the integrand is approximately constant. (The approximation works best for $q \approx N\overline{\ell}$ where the integral is largest.) This leads to

$$p_i \approx \sum_q \exp\left[-\frac{\left(q - N\overline{\ell}\right)^2}{2N\sigma^2}\right] \frac{1}{\sqrt{2\pi N\sigma^2}} \left[\ln(i+1) - \ln i\right] \propto \ln\left(1 + \frac{1}{i}\right),$$

where we have ignored the constants of proportionality which come from the sum over q. We thus find that the distribution of the first digit is *not uniform*, and the properly normalized proportions of $\ln(1+1/i)$ indeed reproduce the probabilities $p_1, \cdots, p_9$ of 0.301, .176, .125, .097, .079, .067, .058, .051, .046 according to Benford's law. (For further information check http://www.treasure-troves.com/math/BenfordsLaw.html.)

****