Massachusetts Institute of Technology

Department of Physics

Physics 8.04 October 11, 1995

Quantum Physics I, Spring 1995

Supplemental Notes on Statistical Averages

1 Introduction

The experimental evidence (G.I. Taylor, Davisson-Germer and others) we have considered indicates that natural phenomena take place as a series of many small discrete events occurring randomly but at a well defined overall random rate described by the square of a quantum probability amplitude. Quantum theory is concerned with the proper description of phenomena and thus intimately involves statistical averages. A full review of statistics is beyond the scope of this course. Fortunately, for our purposes, we need concern ourselves only with the most basic concepts of *the random variable, average, variance/standard deviation and correlation/independence.*

The results we shall derive in sections (2–5) are summarized in the table below. If you are comfortable with these formulas, feel free to skip the corresponding sections. Be sure to read section (6), however, where we apply these results to the spreading of a wave packet propagating in free space and to prove the result quoted in class that if one expects N photons in a given time interval, that the fluctuations in the actual number of photons received will be $\sqrt{N}$. Also, we have some nice plots of probability distributions and their mathematically defined widths in section (4).

TABLE I: BASIC STATISTICAL RESULTS REVIEWED IN THIS NOTE

Result	Conditions	Section
$P_X(x) = \int P_{XY}(x,y)\,dy$	none	3
$< X+Y > = < X > + < Y >$	none	3
$< \alpha X > = \alpha < X >$	none	3
$\Delta X \equiv (\mathrm{Var}X)^{1/2}$ measures width	none	4
$\mathrm{Var}X \equiv < (X - < X >)^2 > = < X >^2 - < X^2 >$	none	4
$\mathrm{Var}\,\alpha X \equiv \alpha^2\,\mathrm{Var}X$	none	4
$P_{XY}(x,y) = P_X(x)\,P_Y(y)$	X and Y independent	5
$< XY > = < X >< Y >$	X and Y independent	5
$\mathrm{Var}(X+Y) = \mathrm{Var}X + \mathrm{Var}\ \mathrm{Y}$	X and Y independent	5
$\Delta(X+Y) = \sqrt{(\Delta X)^2 + (\Delta Y)^2}$	X and Y independent	5
$\Delta N_T = \sqrt{< N_T >}$	Counting Independent Events	6

2 Random Variables

Mathematically, a random variable is defined as a quantity which may take a range of values with a given probability associated with each allowed value. This is closely related to the physical picture we have developed of an *experiment*. In G.I. Taylor's two slit experiment, where a single photon passes through the apparatus at a time, the arrival of each photon gives a separate *measurement*. The *quantity* which we measure in this case is the x coordinate of the spot indicating the location of the photon's arrival on the photographic plate. (See Figure 1.)

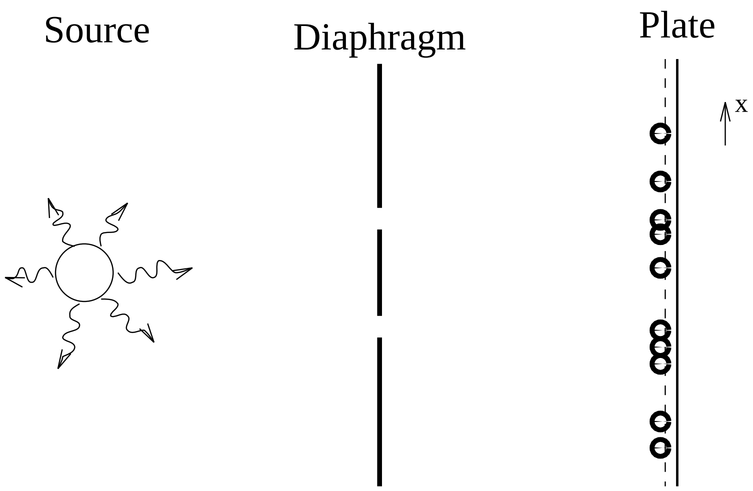

Figure 1: Sketch of G.I. Taylor's Two Slit Apparatus

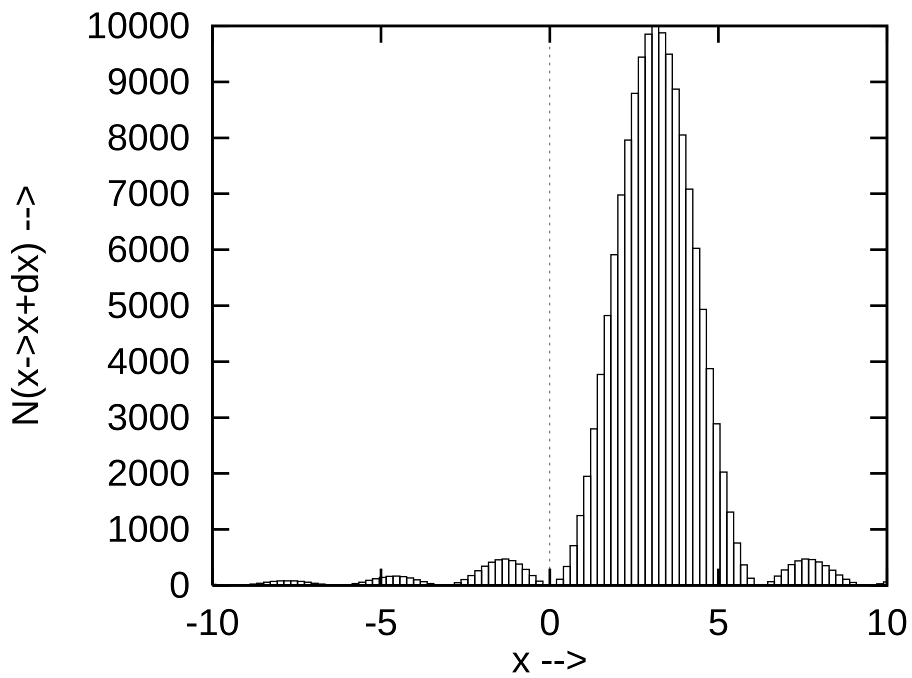

Figure 2: Typical Results from a Taylor Two Slit Experiment

We customarily summarize the results of such an experiment in plot such as Figure 2. This plot tells how many measurements yielded values of x within each of the little bins (ranges in x) indicated in the plot.

We may idealize the situation to where the bins very narrow and we wait long enough to record many photons in each bin. Then by normalizing the results by the total number of measurements (N_{exp}) made in the experiment (the total number of photons recorded) so that $N = N_{exp}P(x)\,dx$ gives the number of photons which arrived in the range $(x \longrightarrow x + dx)$, we determine a probability distribution which characterizes the *experiment*. (See Figure 3.)

Mathematically, we describe the experiment concisely with a random variable X. We say that X is a random variable associated with the position of the recording of the photons. X may take on any value along the photographic plate (which for now we will imagine to extend to $\pm\infty$) and the probability that it falls in the range $(x \longrightarrow x + dx)$ is just $P(x)\,dx$. X thus describes the entire experiment whereas x_i just represents the arrival location of the photon in the experiment.

Note that it is also possible to have a random variable which takes on only discrete values. In this case, we might have an experiment that yields the values $1, 2, ...$ We would then say that the random variable M takes the values $m = 1, 2, ...$ with probabilities $P_m = P_1, P_2....$ All of the results we will demonstrate here are valid for both continuous and discrete random variables. The proofs for the discrete case involve the straight-forward replacement of integrations with summations. We leave these proofs as an exercise to the student.

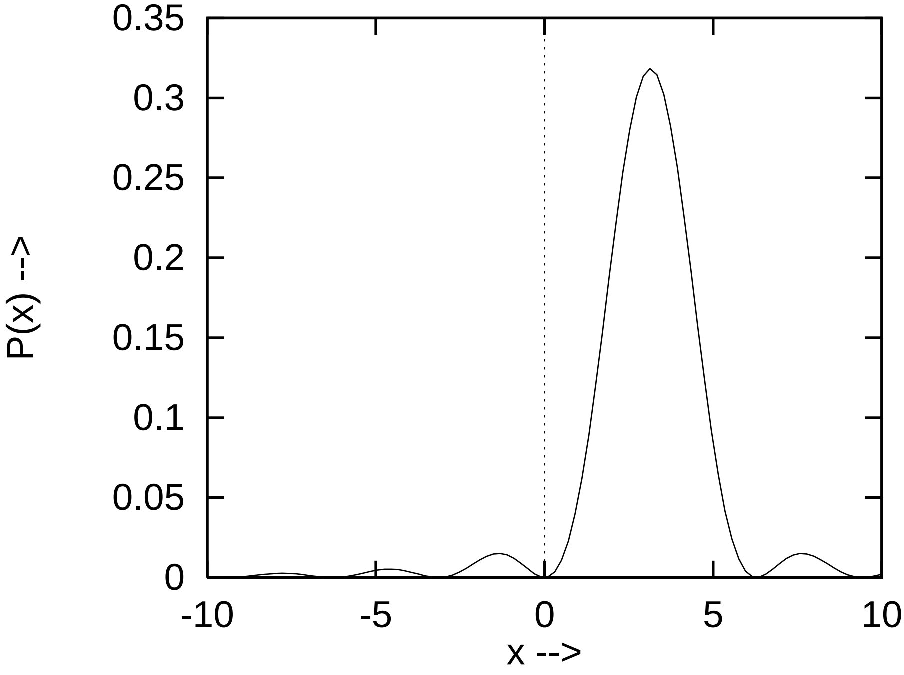

FIGURE 3: PROBABILITY DISTRIBUTION DERIVED FROM THE MEASUREMENTS

3 Averages and Joint Distributions

3.1 Definition of Average

If we now want to compute the average value of x, we can proceed either in the traditionally way of imagining a whole series of *measurements* of x, $\{x_i\}_{i=1}^{N_{exp}}$ and computing the average as

$$\overline{x} \equiv \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} x_i, \tag{1}$$

or, we may get the same result by using the random variable X,

$$< X > \equiv \int_{-\infty}^{\infty} x \cdot P(x)\, dx. \tag{2}$$

The key difference between (1) and (2) is the conceptual focus. In (1) the sum is over the different measurements made; this forces us to think of each individual *measurement*. But in (2), the sum is over the values of the possible outcomes of the experiment and uses the probability distribution explicitly; this frees us to think in terms of the *experiment* as a whole, as a series of potential outcomes occurring with different probabilities. The form (2) is much closer to nature and more convenient for our purposes. We will use it throughout this course. In this note, we will give proofs both in the more familiar form (1) and the newer form (2) to help orient you to it.

The central theorem regarding the averages of random variables is that *the average of a sum of random variables is the sum of their averages*. This is always true regardless of any correlation (which we will describe in detail in section (5)) which may exist between the variables. In the individual measurement picture, this is easy to see.

$$\overline{x+y} = \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} x_i + y_i = \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} x_i + \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} y_i = \overline{x} + \overline{y}.$$

3.2 Joint Probability Distribution

To show this in the random variable picture we need to consider the new concept of *joint probability distribution*. If we had an experiment where each measurement yielded two values x_i and y_i, we would define the joint probability distribution P_{XY} of the random variables X and Y so that the probability of X being in the range $(x \longrightarrow x + dx)$ *and* Y being the range $(y \longrightarrow y + dy)$ is just $P_{XY}(x, y)\, dx\, dy$.

We derive the individual probability distributions $P_X(x)$ (or $P_Y(y)$) from $P_{XY}(x, y)$ through the following argument. $P_(x)\, dx$ is the probability of finding X in the range $(x \longrightarrow x + dx)$ and Y with *any* value. This probability is just $P_X(x)\, dx = \int P_{XY}(x, y)\, dx\, dy$, and so we may take $P_X(x) = \int P_{XY}(x, y)\, dy$. A similar argument gives $P_Y(y) = \int P_{XY}(x, y)\, dx$.

3.3 Basic Properties of the Average

With this established, we then have for the average of the sum of two random variables,

$$\begin{aligned}
< X + Y > \equiv \int\int (x+y) \cdot P_{XY}(x,y)\, dxdy &= \int\int x P_{XY}(x,y)\, dxdy + \int\int y P_{XY}(x,y)\, dxdy \\
&= \int x \left[\int P_{XY}(x,y)\, dy\right] dx + \int y \left[\int P_{XY}(x,y)\, dx\right] dy \\
&= \int x P_X(x)\, dx + \int y P_Y(y)\, dy \\
&= < X > + < Y > .
\end{aligned}$$

For the mathematically minded, the general case of any finite sum of random variable then follows from induction.

If instead of add two random variables, we multiply a single random variable by a constant, we find,

$$\overline{\alpha x} = \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} \alpha x_i = \alpha \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} x_i = \alpha \overline{x},$$

$$< \alpha X >= \int \alpha x \cdot P(x)\, dx = \alpha \int x \cdot P(x)\, dx = \alpha < X > .$$

4 Variance and Standard Deviation

4.1 Definition of Variance

While an experiment may yield well-defined *average* results, generally every measurement does not yield precisely the same result and there is some spread or variation in the results (as in Figure 3 above). The mathematical concept of *variance* is an attempt to give a formal measure of this spread.

A first, failed, attempt at defining the variance might be to compute the average deviation $(x - \overline{x})$ or $(X- < X >)$ from the average. But, one then quickly finds

$$V = \overline{x - \overline{x}} = \overline{x} - \overline{\overline{x}} = \overline{x} - \overline{x} = 0,$$

$$V =< (X- < X >) >=< X > - << X >>=< X > - < X >= 0.$$

Here we have used the fact that the average value of a constant is just that constant. For instance

$$\overline{\overline{x}} = \frac{1}{N_{exp}} \sum_{i=1}^{N_{exp}} \overline{x} = \frac{1}{N_{exp}} \cdot N_{exp} \cdot \overline{x} = \overline{x},$$

$$<< X >>= \int < X > \cdot P(x)\, dx =< X > \int P(X)\, dx =< X > \cdot 1 =< X > .$$

Clearly, this average deviation is *not* a good measure of the width of the distribution in X. The problem is that we get both positive and *negative* contributions that then cancel and give us zero. To prevent this from happening, we can take the average of the *square* of the deviation from the average. This is the proper *definition* of the variance of X, which is written either as σ_x^2 in the individual measurement picture or as $\mathrm{Var}\, X$ in the random variable picture.

$$\begin{aligned}
\sigma_x^2 \equiv \overline{(x-\overline{x})^2} = \overline{x^2 - 2x\overline{x} + \overline{x}^2} &= \overline{x^2} - \overline{2x\overline{x}} + \overline{\overline{x}^2} \\
&= \overline{x^2} - 2\,\overline{x}\,\overline{x} + \overline{x}^2 \\
\sigma_x^2 &= \overline{x^2} - \overline{x}^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\, X \equiv < (X- < X >)^2 > &= < X^2 - 2X < X > + < X >^2 > \\
&= < X^2 > -2 < X >< X > + < X >^2 \\
\mathrm{Var}\, X &= < X^2 > - < X >^2 .
\end{aligned}$$

We have used the facts we established in section (3) that the average of a sum is the sum of the averages and that the average of a constant times a random variable is just the constant times the average of the random variable.

4.2 Basic Properties of the Variance

One useful result about variances which is relatively easy to show is that because the variance gives a measure or the *square* of the width of a distribution, *the variance of a constant times a random variable is the* square *of the constant times the variance of the random variable.*

$$\begin{aligned}
\sigma_{\alpha x}^2 &\equiv \overline{(\alpha x)^2} - \overline{\alpha x}^2 \\
&= \overline{\alpha^2 x^2} - (\alpha \overline{x})^2 \\
&= \alpha^2 \overline{x^2} - \alpha^2 \overline{x}^2 \\
&= \alpha^2 \left(\overline{x^2} - \overline{x}^2 \right) \\
\sigma_{\alpha x}^2 &= \alpha^2 \sigma_x^2
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}\, \alpha X &\equiv < (\alpha X)^2 > - < \alpha X >^2 \\
&= < \alpha^2 X^2 > - (\alpha < X >)^2 \\
&= \alpha^2 < X^2 > -\alpha^2 < X >^2 \\
&= \alpha^2 \left(< X^2 > - < X >^2 \right) \\
\mathrm{Var}\, \alpha X &= \alpha^2 \mathrm{Var}\, X
\end{aligned}$$

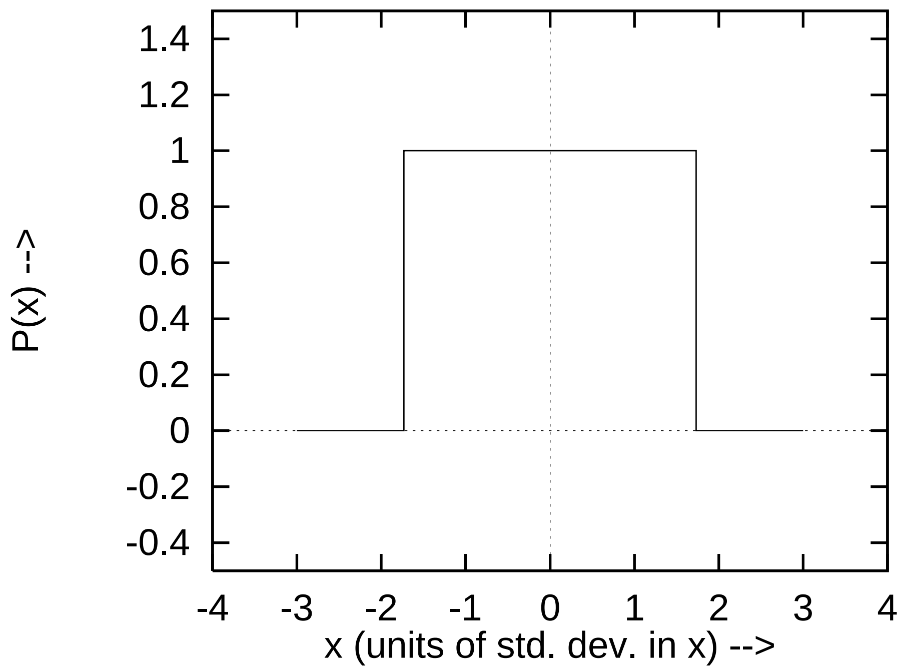

FIGURE 4: UNIFORM PROBABILITY DISTRIBUTION FOR X

4.3 Examples of Variance

At this point, it is useful to try some examples. First, we may have a uniform probability distribution localized in the range $(-\Delta/2 \longrightarrow \Delta/2)$,

$$P(X) = \begin{cases} 0 & \text{if } |X| > \Delta/2 \\ 1/\Delta & \text{if } |X| < \Delta/2 \end{cases}$$

This distribution is sketched in figure 4. (The units in this figure are not units of Δ but are units of the standard deviation, which we will show how to compute presently.)

Note that we have taken proper care that the distribution is normalized. X must have some value, so

$$< 1 > \equiv \int P(x)\, dx = 1,$$

which is true in this case. Clearly also, $< X > = \int x \cdot P(x)\, dx = \frac{1}{\Delta}\int_{-\Delta/2}^{\Delta/2} x\, dx = 0$. We thus have,

$$\begin{aligned} \operatorname{Var} X &\equiv < X^2 > - < X >^2 \\ &= < X^2 > \\ &\equiv \int x^2 \cdot P(x)\, dx \\ &= \int_{-\Delta/2}^{\Delta/2} x^2 \cdot \frac{1}{\Delta}\, dx \end{aligned}$$

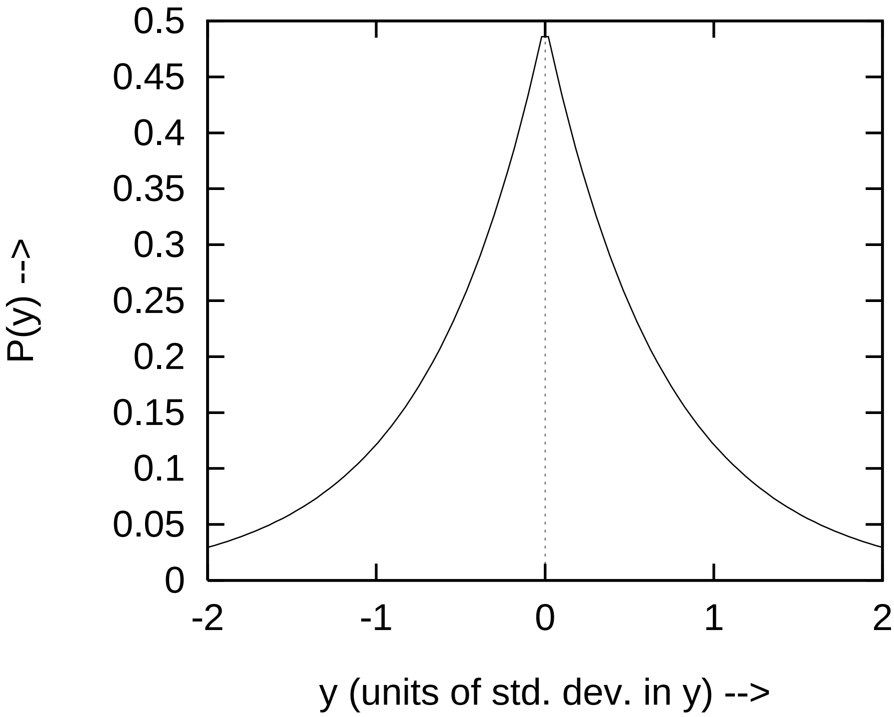

Figure 5: Exponential Probability Distribution for Y

$$
\begin{aligned}
&= 2 \cdot \frac{1}{3} \left(\frac{\Delta}{2}\right)^3 \frac{1}{\Delta} \\
&= \frac{\Delta^2}{12}.
\end{aligned}
$$

$\mathrm{Var}\, X$ is now non-zero and clearly increases with the spread in the distribution Δ.

We could also have a random variable Y distributed according to an exponential distribution

$$P(y) = \frac{1}{2\Delta} e^{-|y|/\Delta},$$

as in Figure 5.

This distribution is also properly normalized ($< 1 >= 1$) with zero average ($< Y >= 0$). Now,

$$
\begin{aligned}
\mathrm{Var}\, Y &\equiv < Y^2 > - < Y >^2 \\
&= < Y^2 > \\
&\equiv \int y^2 \cdot P(y)\, dy \\
&= \int_{-\infty}^{\infty} y^2 \cdot \frac{1}{2\Delta} e^{-|y|/\Delta}\, dy \\
&= \frac{1}{2} \cdot 2 \cdot \int_0^{\infty} \Delta^2 \left(\frac{y}{\Delta}\right)^2 e^{-\frac{y}{\Delta}}\, d\left(\frac{y}{\Delta}\right) \\
&= \Delta^2 \int_0^{\infty} u^2 e^{-u}\, du \\
&= \Delta^2 (2!)
\end{aligned}
$$

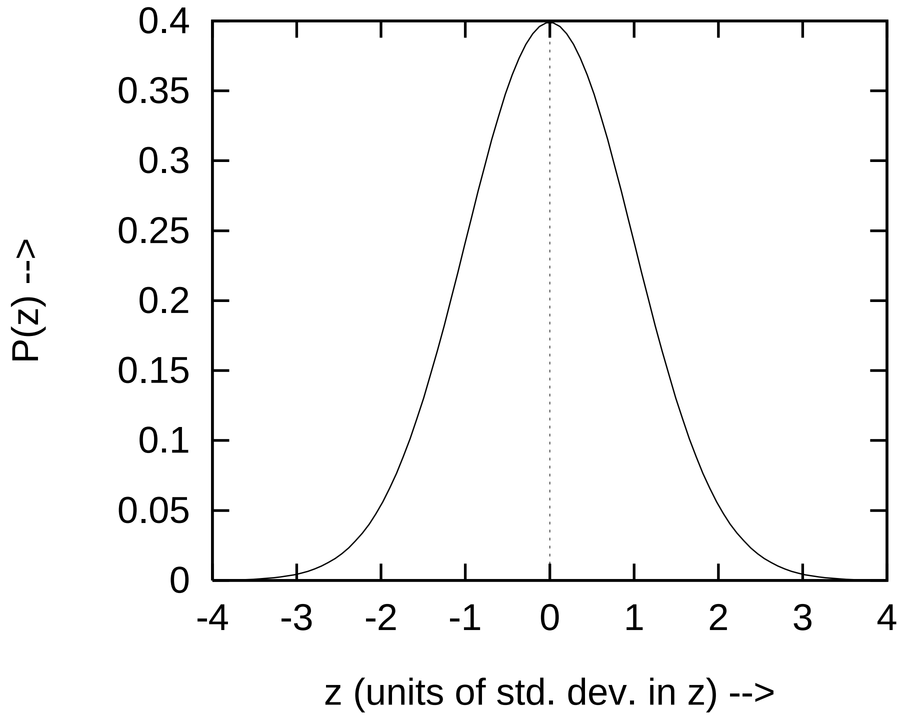

FIGURE 6: NORMAL PROBABILITY DISTRIBUTION FOR Z

$$= \quad 2\Delta^2,$$

which again is non-zero and grows with the width of the distribution.

Finally, we may have a random variable Z distributed according to the normal distribution,

$$P(z) = \frac{1}{\sqrt{2\pi}\Delta} e^{-\frac{1}{2}\left(\frac{z}{\Delta}\right)^2},$$

as in Figure 6.

Here,

$$\begin{aligned}
\mathrm{Var}\, Z &\equiv\ < Z^2 > - < Z >^2 \\
&=\ < Z^2 > \\
&\equiv \int z^2 \cdot P(z)\, dz \\
&= \int_{-\infty}^{\infty} z^2 \cdot \frac{1}{\sqrt{2\pi}\Delta} e^{-\frac{1}{2}\left(\frac{z}{\Delta}\right)^2} \\
&= \frac{\Delta^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \left(\frac{z}{\Delta}\right)^2 e^{-\frac{1}{2}\left(\frac{z}{\Delta}\right)^2} d\left(\frac{z}{\Delta}\right) \\
&= \frac{\Delta^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} u^2 e^{-u^2/2}\, du \\
&= \frac{\Delta^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} -u\, de^{-u^2/2} \\
&= \frac{\Delta^2}{\sqrt{2\pi}} \int_{-\infty}^{\infty} e^{-u^2/2}
\end{aligned}$$

$$
\begin{aligned}
&= \frac{\Delta^2}{\sqrt{2\pi}} \cdot \sqrt{2\pi} \\
&= \Delta^2.
\end{aligned}
$$

In all three of these cases we found that the variance measures the *square* of the parameter controlling the width Δ of the distribution. This leads us to our final definition, that of the *standard deviation*, which is meant to give a *direct* measure of the width of a distribution. We define the standard deviation as the square root of the variance,

$$
\sigma_x \equiv \sqrt{\sigma_x^2} = \sqrt{\overline{(x - \overline{x})^2}},
$$

$$
\Delta X \equiv \sqrt{\mathrm{Var}\, X} = \sqrt{< (X - < X >)^2 >}.
$$

Because of its form, the standard deviation is sometimes referred to as the root mean square (rms) deviation. This is the unit that was used along the horizontal axis of the plots for the distributions of X, Y, and Z in Figures 4-6. From those figures you can see that the standard deviation does correspond to what we would qualitatively associate with the approximate width of these distributions.

The table below summarizes our results for these three distributions.

TABLE II: WIDTHS FOR THREE BASIC DISTRIBUTIONS

Variable	Distribution	Standard Deviation
X	Uniform, $\pm\Delta/2$	$\Delta X = \Delta/\sqrt{12}$
Y	Exponential Decay, $1/e$ at Δ	$\Delta Y = \sqrt{2}\Delta$
Z	Normal Distribution, std. dev. Δ	$\Delta Z = \Delta$

5 Independence and the Variance of a Sum of Independent Variables

One very useful property of the variance is that *the variance of the sum of* independently *distributed random variables is the sum of the variances of the individual random variables.* It is important to note that this is true only if the random variables are independent and *un*correlated.

5.1 Independence of Random Variables

If X and Y are two random variables, we say that they are independent if we get the same distribution for X whether we plot all of our results for X as below,

or whether we plot our results for X when Y falls within a small range of any of its possible values.

From section (3), we know that the probability that X is in the range $(x \longrightarrow x + dx)$ given that Y is in the range $(y \longrightarrow y + dy)$ is just $P_{XY}(x, y)\, dx\, dy$. The fact that all of the distribution curves generated for X, even when Y takes on different specific values, look precisely the same after proper normalization by some y-dependent factor $N(y)$ means that

$$
P_x(x)\, dx\, N(y) \quad = \quad P_{XY}(x, y)\, dx\, dy.
$$

To find $N(y)$, we integrate both sides with respect to x,

$$
\begin{aligned}
\left(\int P_X(x)\, dx\right) N(y) &= \left(\int P_{XY}(x, y)\, dx\right) dy \\
1 \cdot N(y) &= P_Y(y)\, dy \\
N(y) &= P_Y(y)\, dy,
\end{aligned}
$$

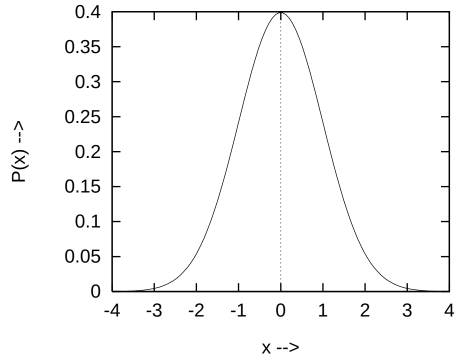

Figure 7a: Tabulating All Results for X

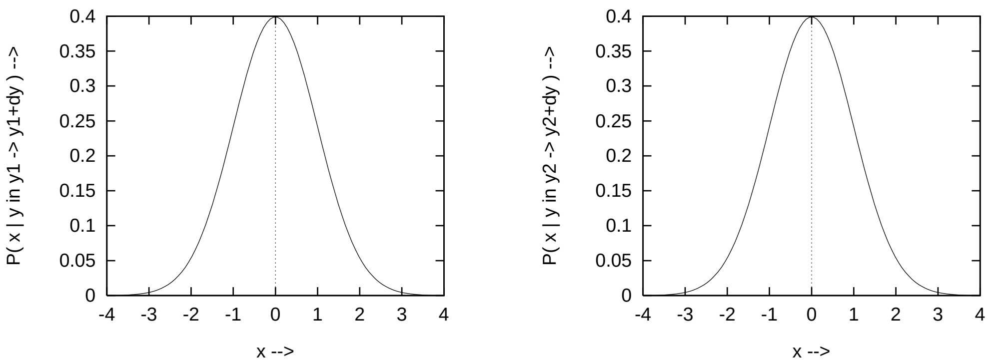

Figure 7b-c: Tabulating Results for X with Y in the Ranges $(y_1 \longrightarrow y_1 + dy)$ or $(y_2 \longrightarrow y_w + dy)$.

and so,

$$\begin{aligned} P_{XY}(x,y)\,dx\,dy &= P_X(x)\,dx\;P_Y(y)\,dy \\ P_{XY}(x,y) &= P_X(x)P_Y(y). \end{aligned}$$

This is the mathematical expression of our physical notion of the independence of experimental variables. *If two random variables are* independent*, then their joint probability distribution is just the* product *of their individual probability distributions.*

5.2 Average of the Product of Two Independent Random Variables

Armed with definition, we now have an immediate and useful *lemma. The average of the product of two independently distributed random variables is the product of their individual averages.*

$$\begin{aligned} <XY> &= \int xy\cdot P_{XY}(x,y)\,dx\,dy \\ &= \int xyP_X(x)\,P_Y(y)\,dx\,dy \\ &= \left(\int xP_X(x)dx\right)\left(\int yP_Y(y)dy\right) \\ <XY> &= <X><Y>. \end{aligned}$$

5.3 Variance of the Sum of Two Independent Random Variables

Our main result for variances now follows directly,

$$\begin{aligned} \sigma^2_{x+y} &\equiv \overline{(x+y)^2-(\overline{x+y})^2} \\ &= \overline{(x^2+2xy+y^2)-(\overline{x}^2+2\,\overline{x}\,\overline{y}+\overline{y}^2)} \\ &= \left(\overline{x^2}-\overline{x}^2\right)+\left(\overline{y^2}-\overline{y}^2\right)+2\left(\overline{xy}-\overline{x}\,\overline{y}\right) \\ &= \sigma^2_x+\sigma^2_y+2\cdot 0 \\ \sigma^2_{x+y} &= \sigma^2_x+\sigma^2_y \end{aligned}$$

$$\begin{aligned} \mathrm{Var}\,(X+Y) &\equiv <(X+Y)^2-<X+Y>^2> \\ &= <X^2+2XY+Y^2>-(<X>+<Y>)^2) \\ &= <X>^2+2<XY>+<Y>^2-\left(<X>^2+2<X><Y>+<Y>^2\right) \\ &= \left(<X>^2-<X>^2\right)+\left(<Y>^2-<Y>^2\right)+2\left(<XY>-<X><Y>\right) \\ &= \mathrm{Var}\,X+\mathrm{Var}\,Y+2\cdot 0 \\ \mathrm{Var}\,(X+Y) &= \mathrm{Var}\,X+\mathrm{Var}\,Y \end{aligned}$$

The fact that the *variances* of independent variables add algebraically gives us a nice mnemonic for the addition of standard deviations. Because variances, the square of standard deviations add algebraically, the standard deviations add geometrically like the sides of a right triangle.

$$\begin{aligned} \sigma_{x+y} &= \sqrt{\sigma^2_x+\sigma^2_y} \\ \Delta(X+Y) &= \sqrt{(\Delta X)^2+(\Delta Y)^2} \end{aligned}$$

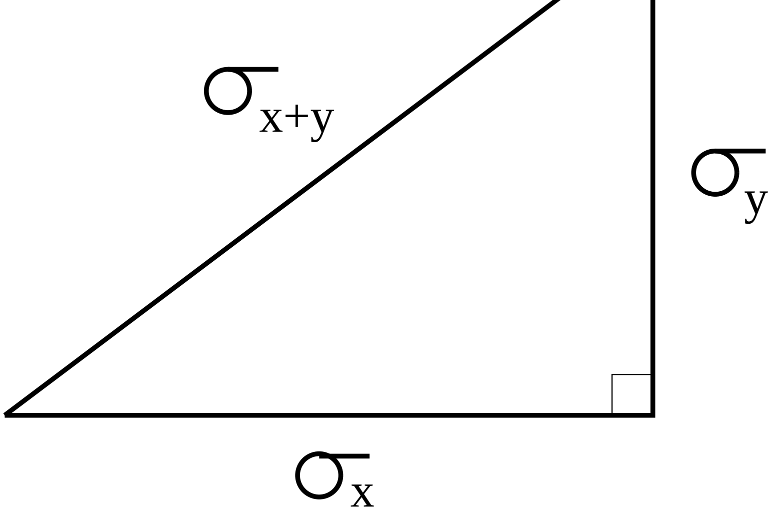

Figure 8: Addition of Standard Deviations

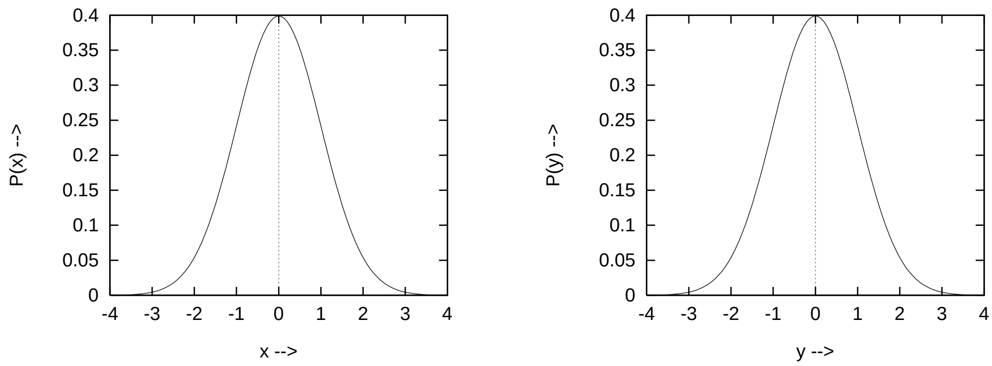

Figure 9a-b: Variant Distributions of X and Y in a Correlated Example

5.4 The Variance of the Sum of Correlated Variables May Be Zero!

Note that these results need not hold if X and Y are not independent, are *correlated.* As an instructive example, note that it is possible to design an experiment where X and Y are distributed with significant variances as in Figures 9a-b,

but where there is no variance in their sum whatsoever, ($\sigma_{x+y} = 0$)!

An example of an apparatus generating these results would be a random number pair generator than just generates an internal random number Z with a significant variance but then reports to its two outputs $X = -Z$ and $Y = Z$. (See Figure 10.)

While X and Y will have the same spread in their distributions as Z, the sum $X+Y$ will always be exactly zero and exhibit no spread at all. This is related to the observation that while there is a spread when measuring the momentum p or the position r of the electron in the ground state of a hydrogen atom, measuring the energy $E = \frac{p^2}{2m} - \frac{e^2}{r}$, always gives precisely the same result of $-13.6\,eV$.

6 Applications

6.1 Minimal Spreading of a Wave Packet

Suppose we have an experiment which produces a probability wave packet describing a particle of mass m in free space. At time $t = 0$, the packer has a width ΔX centered about $< X >$ and travels at a velocity $< V >$. By the uncertainty principle, we know that there is an uncertainty in the momentum of the packet exceeding $\Delta P \geq \frac{\hbar}{2\Delta X}$. Let us suppose that we begin with a packet with the absolute minimum uncertainty. (We will be precise here with the uncertainty principle because we are now using precise mathematical definitions for ΔX and ΔP and because we later will be able to compare our result here with a calculation using the full formal wave theory.) Associated with the uncertainty in momentum, there is also an uncertainty in the velocity of the particle, $\Delta V = \Delta P/m = \hbar/(2m\Delta X)$.

Because we are in free space and there is no force to change the velocity of the particle, if we wait a time $t > 0$ the velocity of the particle will not change

$$V(t) = V(0) \equiv V$$

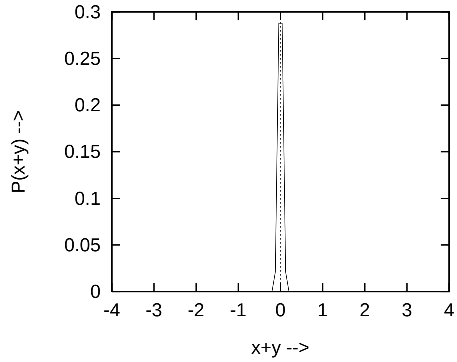

Figure 9c: Invariant Distribution of $X + Y$ in the Correlated Example

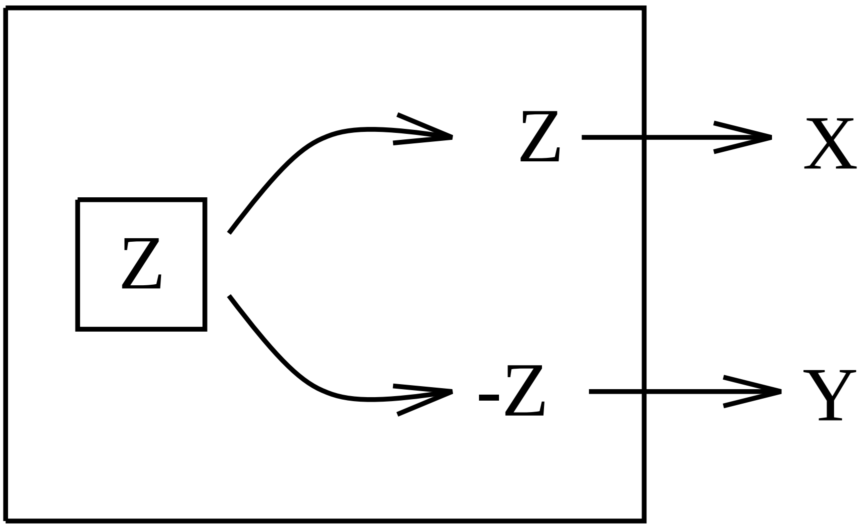

Figure 10: Random Number Pair Generator Producing Results from Figure 9

and its new position will be given by the usual formula

$$X(t) = X(0) + Vt = X + Vt.$$

The uncertainty (standard deviation) in $X(t)$ now comes from uncertainties both in X and V. *Assuming* that X and V are not correlated and are *independent*, we then have

$$\begin{aligned}
(\Delta V(t))^2 \equiv \mathrm{Var}\,V(t) = \mathrm{Var}\,V &\equiv (\Delta V)^2 = \left(\frac{\hbar}{2m\,\Delta X}\right)^2 \\
(\Delta X(t))^2 \equiv \mathrm{Var}\,X(t) = \mathrm{Var}\,(X + Vt) &= \mathrm{Var}\,X + t^2 \mathrm{Var}\,V = \Delta X + \left(\frac{\hbar t}{2m\,\Delta X}\right)^2,
\end{aligned}$$

So that,

$$\begin{aligned}
\Delta V(t) &= \frac{\hbar}{2m\,\Delta X} \\
\Delta X(t) &= \sqrt{(\Delta X)^2 + \left(\frac{\hbar t}{2m\,\Delta X}\right)^2}
\end{aligned}$$

Note that $\Delta V(t)$ *does not change with time* as the packet spreads because there is no force acting to change the velocity of the particle in each measurement. This does not violate the uncertainty principle because $\Delta X\,\Delta P = \Delta X\, m\Delta V$, which grows with time, continues to exceed $\hbar/2$. The uncertainty principle is an *in*equality, not an equality!

6.2 Poisson(Counting) Statistics

The picture we have of an experiment involves two phases, (1) a probability wave propagates through the apparatus, (2) particles arrive according to the distribution given by the square of the amplitude of the wave. If we repeat our experiment many times in a unit time interval, as in a two-slit Young interference experiment where the source sends many particles per second through the apparatus, then the probability distribution (as in Figure 3) gives us the average rate of arrival of photons in each region of the plate. If we wait a sampling time T and the source emits photons at a rate of I, then the *average* number of particles N_T we expect to arrive on the plate with positions in the interval $(x \longrightarrow x + dx)$ is

$$< N_T >= ITP(x)dx = RT.$$

This is just the sampling time T times some average rate R determined by the distribution $P(x)$. Because of the randomness inherent in the process, however, if we were to repeat the experiment of counting photons arriving in the interval $(x \longrightarrow x + dx)$ during intervals of time T many times, we would not find the same number of photons N_T in each experiment. There will be a distribution in the values of N_T as in Figure 11 below. N_T is a new random variable, with an average value $< N_T >$ and a variance of its own. Using very general arguments we will now demonstrate the result quoted in class

$$\Delta N_T \equiv \sqrt{\mathrm{Var}\,N_T} = \sqrt{< N_T >},$$

Our arguments are very general and will apply to nearly all counting experiments. The distribution we are discussion is so important that it bears its own name, the *Poisson distribution.*

We begin by showing that $\mathrm{Var}\,X_T \propto T$, from which $\Delta N_T \equiv \sqrt{\mathrm{Var}\,X_T} \propto T^{1/2} \propto < N_T >^{1/2}$ will follow quickly. To show $\mathrm{Var}\,X_T \propto T$ we note use the facts that the variance of independent random variables add and that the photons in the experiment arrive randomly or *independently* of one another.

This means that if we were, for instance, to divide our counting time interval from $(t = 0 \longrightarrow t = T)$ into two separate intervals, interval 1 from $(t = 0 \longrightarrow t = T/2)$ and interval 2 from $(t = T/2 \longrightarrow t = T)$, then we could define two *independently distributed* random variables $N^{(1)}_{T/2}$ and

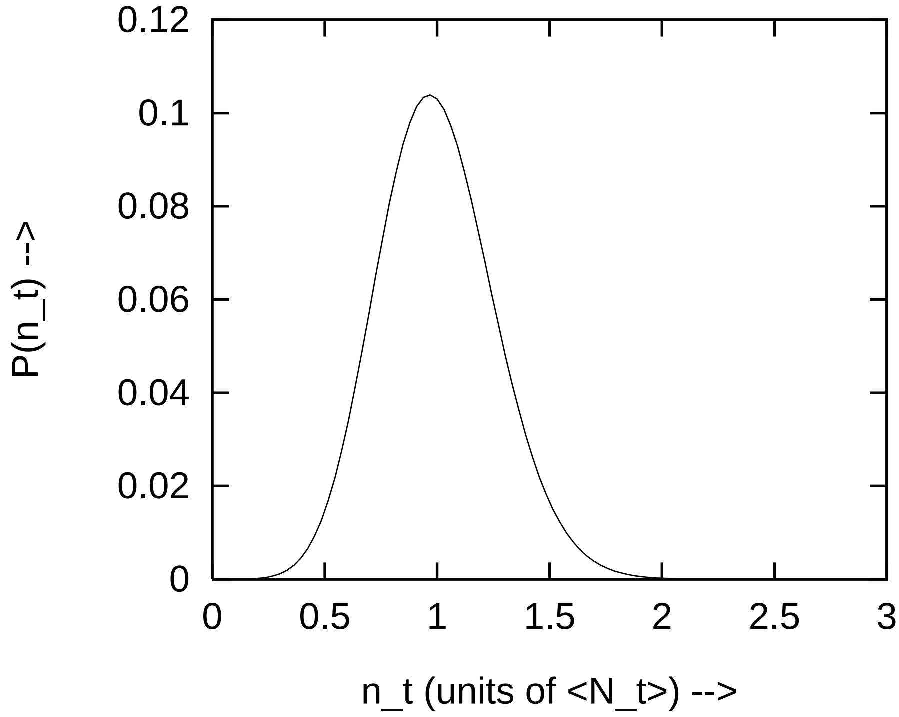

FIGURE 11: POISSON DISTRIBUTION

$N_{T/2}^{(2)}$, which each represent a counting experiment of length $T/2$ but carried out at two different times. The sum of these two random variables is just the total number of particles arriving in our original time interval, so that $N_T = N_{T/2}^{(1)} + N_{T/2}^{(2)}$. Finally, we know that as long as the source intensity is constant, the counting statistics do not depend on *when* we do our counting; they depend only on *how long* we do the counting. Thus, we may take $\mathrm{Var}\, N_{T/2}^{(1)} = \mathrm{Var}\, N_{T/2}^{(2)} = \mathrm{Var}\, N_{T/2}$, where $\mathrm{Var}\, N_{T/2}$ is the variance of any counting experiment over a length of time $T/2$ at any point in time. We have then,

$$\begin{aligned}
\mathrm{Var}\, N_T &= \mathrm{Var}\,(N_{T/2}^{(1)} + N_{T/2}^{(2)}) \\
&= \mathrm{Var}\,(N_{T/2}^{(1)}) + \mathrm{Var}\,(N_{T/2}^{(2)}) \\
&= \mathrm{Var}\, N_{T/2} + \mathrm{Var}\, N_{T/2} \\
&= 2\,\mathrm{Var}\, N_{T/2}
\end{aligned}$$

We could have, instead of just two divisions, divided our time interval T into many small time intervals of length δ. Our total time T is then divided into T/δ intervals and we find,

$$\begin{aligned}
\mathrm{Var}\, N_T &= \mathrm{Var}\,(N_\delta^{(1)} + N_\delta^{(2)} + \ldots + N_\delta^{(T/\delta)}) \\
&= \mathrm{Var}\, N_\delta^{(1)} + \mathrm{Var}\, N_\delta^{(2)} + \ldots + \mathrm{Var}\, N_\delta^{(T/\delta)} \\
&= \frac{T}{\delta}\mathrm{Var}\, N_\delta \\
\mathrm{Var}\, N_T &= \left(\frac{\mathrm{Var}\, N_\delta}{\delta}\right) T
\end{aligned} \tag{3}$$

We thus have shown that $\mathrm{Var}\, N_T \propto T$ by just using the fact that the variances of *independent* random variables add and by dividing the counting interval into a set of small time intervals whose number varies in proportion to the sampling time interval.

We now need only evaluate the proportionality constant $\left(\frac{\mathrm{Var}\, N_\delta}{\delta}\right)$. This is most easily done as the time interval δ becomes very small. As $\delta \longrightarrow 0$, the probability of counting zero photons P_0 approaches unity, $P_0 \rightarrow 1$; the probability of counting a single photon P_1 is very small $P_1 \propto \delta << 1$; and the probability of counting two or more photons is negligible, $P_n \propto \delta^n, n \geq 2$. We can now determine P_1 from the average rate. Counting over a time δ, on average we expect $\delta \cdot R$ photons and so,

$$\delta \cdot R = < N_\delta > \equiv 0 \cdot P_0 + 1 \cdot P_1 + 2 \cdot P_2 + \ldots + \simeq P_1.$$

The variance is then

$$\begin{aligned}
\mathrm{Var}\, N_\delta &\equiv < N_\delta^2 > - < N_\delta >^2 \\
&= 0^2 \cdot P_0 + 1^2 \cdot P_1 + 2^2 \cdot P_2 + \ldots - < N_\delta >^2 \\
&\simeq P_1 - P_1^2 \\
&\simeq \delta \cdot R,
\end{aligned}$$

where we ignore terms δ^2 and higher. We now have our proportionality constant,

$$\lim_{\delta \longrightarrow 0} \frac{\mathrm{Var}\, N_\delta}{\delta} = R.$$

Combining this with (3), we have our final result,

$$(\Delta N_T)^2 = \mathrm{Var}\, N_T = \left(\lim_{\delta \longrightarrow 0} \frac{\mathrm{Var}\, N_\delta}{\delta}\right) T = RT = < N_T >,$$

and so

$$\Delta N_T = \sqrt{< N_T >}$$

Note that this result is very general and not restricted to photons. Our derivation depends only on the counting of independent random events and so could apply equally well to counting either the arrival of electrons, the decay of nuclei or the arrival of cars in the parking lot at the shopping mall. On Problem Set 2, you showed how to apply this result to estimate the value of Planck's constant from observations you can make in your own room at night.