In a heuristic sense, the assumptions of the statistical model appear to be valid if the signal is sufficiently complex and the quantization steps are sufficiently small, so that the amplitude of the signal is likely to traverse many quantization steps from sample to sample.
Review: Assumptions about e[n]

- e[n] is a sample sequence of a stationary random process.
- e[n] is uncorrelated with the sequence x[n].
- The random variables of the error process e[n] are uncorrelated; i.e., the error is a white-noise process.
- The probability distribution of the error process is uniform over the range of quantization error (i.e., without being clipped).
- The assumptions would not be justified. However, when the signal is a complicated signal (such as speech or music), the assumptions are more realistic.
  - Experiments have shown that, as the signal becomes more complicated, the measured correlation between the signal and the quantization error decreases, and the error also becomes uncorrelated.
Review: Quantization error analysis

\[-\frac{\Delta}{2} < e[n] \leq \frac{\Delta}{2}\]

- $e[n]$ is a white noise sequence. The probability density function of $e[n]$ is a uniform distribution
The mean value of $e[n]$ is zero, and its variance is

$$\sigma_e^2 = \int_{-\Delta/2}^{\Delta/2} e^2 \frac{1}{\Delta} de = \frac{\Delta^2}{12}$$

Since $\Delta = \frac{X_m}{2^B}$

For a $(B+1)$-bit quantizer with full-scale value $X_m$, the noise variance, or power, is

$$\sigma_e^2 = \frac{2^{-2B} X_m^2}{12}$$
Review: Quantization error analysis

- A common measure of the amount of degradation of a signal by additive noise is the signal-to-noise ratio (SNR), defined as the ratio of signal variance (power) to noise variance. Expressed in decibels (dB), the SNR of a (B+1)-bit quantizer is

\[
SNR = 10 \log_{10} \left( \frac{\sigma_x^2}{\sigma_e^2} \right) = 10 \log_{10} \left( \frac{12 \cdot 2^{2B} \sigma_x^2}{X_m^2} \right)
\]

\[
= 6.02B + 10.8 - 20 \log_{10} \left( \frac{X_m}{\sigma_x} \right)
\]

- Hence, the SNR increases approximately 6dB for each bit added to the word length of the quantized samples.
We consider the analog signal $x_a(t)$ as zero-mean, wide-sense-stationary, random process with power-spectral density denoted by $\phi_{x_a x_a}(e^{jw})$ and the autocorrelation function by $\phi_{x_a x_a}(\tau)$.

To simplify our discussion, assume that $x_a(t)$ is already bandlimited to $\Omega_N$, i.e.,

$$\phi_{x_a x_a}(j\Omega) = 0, \quad |\Omega| \geq \Omega_N,$$
We assume that $2\pi/T = 2M\Omega_N$.

M is an integer, called the **oversampling ratio**.

Oversampling

Oversampled A/D conversion with simple quantization and down-sampling

Decimation with ideal low-pass filter
Using the additive noise model, the system can be replaced by

\[ e[n] \]

\[ \hat{x}[n] = x[n] + e[n] \]

\[ x_d[n] = x_{da}[n] + x_{de}[n] \]

Its output \( x_d[n] \) has two components, one due to the signal input \( x_a(t) \) and one due to the quantization noise input \( e[n] \). Denote these components by \( x_{da}[n] \) and \( x_{de}[n] \), respectively.

Decimation with ideal low-pass filter
Goal: determine the ratio of signal power $\mathbb{E}\{x_{da}^2\}$ to the quantization-noise power $\mathbb{E}\{x_{de}^2\}$, $\mathbb{E}\{\cdot\}$ denotes the expectation value.

As $x_a(t)$ is converted into $x[n]$, and then $x_{da}[n]$, we focus on the power of $x[n]$ first.

Let us analysis this in the time domain. Denote $\phi_{xx}[n]$ and $\phi_{xx}[e^{jw}]$ be the autocorrelation and power spectral density of $x[n]$, respectively.

By definition, $\phi_{xx}[m] = \mathbb{E}\{x[n+m]x[n]\}$. 

Signal component (assume $e[n]=0$)
Power of $x[n]$ (assume $e[n]=0$)

- Since $x[n] = x_a(nT)$, it is easy to see that
  \[
  \phi_{xx}[m] = \varepsilon \{x[n + m]x[n]\}
  = \varepsilon \{x_a((n + m)T)x_a(nT)\}
  = \phi_{x_a x_a}(mT)
  \]

- That is, the autocorrelation function of the sequence of samples is a sampled version of the autocorrelation function.

- The wide-sense-stationary assumption implies that $\varepsilon \{x_a^2(t)\}$ is a constant independent of $t$. It then follows that
  \[
  \varepsilon \{x^2[n]\} = \varepsilon \{x_a^2(nT)\} = \varepsilon \{x_a^2(t)\}
  \]
  for all $n$ or $t$. 
Power of $x_{da}[n]$ (assume $e[n]=0$)

- Since the decimation filter is an ideal lowpass filter with cutoff frequency $w_c = \pi/M$, the signal $x[n]$ passes unaltered through the filter.

- Therefore, the downsamsled signal component at the output, $x_{da}[n]=x[nM]=x_a(nMT)$, also has the same power.

- In sum, the above analyses show that

$$\varepsilon \{x_{da}^2[n]\} = \varepsilon \{x^2[n]\} = \varepsilon \{x_a^2(t)\}$$

which shows that the power of the signal component stays the same as it traverse the entire system from the input $x_a(t)$ to the corresponding output component $x_{da}[n]$. 
Power of the noise component

- According to previous studies, let us assume that $e[n]$ is a wide-sense-stationary white-noise process with zero mean and variance

$$\sigma_e^2 = \frac{\Delta^2}{12}$$

- Consequently, the autocorrelation function and power density spectrum for $e[n]$ are,

$$\phi_{ee}[m] = \sigma_e^2 \delta[m]$$

- The power spectral density is the DTFT of the autocorrelation function. So,

$$\phi_{ee}(e^{j\omega}) = \sigma_e^2, \quad -\pi < \omega < \pi$$
Although we have shown that the power in either $x[n]$ or $e[n]$ does not depend on $M$, we will show that the noise component $x_{de}[n]$ does not keep the same noise power. It is because that, as the oversampling ratio $M$ increases, less of the quantization noise spectrum overlaps with the signal spectrum, as shown below.
Illustration of frequency and amplitude scaling

Since oversampled by $M$, the power spectrum of $x_a(t)$ and $x[n]$ in the frequency domain are illustrated as follows.

\[ \Phi_{x_a x_a}(j\Omega) \]

\[ \Phi_{xx}(e^{j\omega}) \]

\[ \frac{1}{T} = \frac{\Omega_N M}{\pi} \]
By considering both the signal and the quantization noise, the power spectra of $x[n]$ and $e[n]$ in the frequency domain are illustrated as:

$$\frac{1}{T} = \frac{\Omega_N M}{\pi}$$

$$\Phi_{xx}(e^{j\omega})$$

$$\Phi_{ee}(e^{j\omega}) = \sigma_e^2$$
Then, by ideal low pass with cutoff $\omega_c = \pi / M$ in the decimation, the noise power at the output becomes

$$E\{e^2[n]\} = \frac{1}{2\pi} \int_{-\pi/M}^{\pi/M} \sigma_e^2 dw = \frac{\sigma_e^2}{M}$$
Next, the lowpass filtered signal is downsampled, and as we have seen, the signal power remains the same. Hence, the power spectrum of $x_{da}[n]$ and $x_{de}[n]$ in the frequency domain are illustrated as follows:

\[ \frac{1}{T'} = \frac{\Omega_N}{\pi} \]

\[ \Phi_{x_{da}x_{da}}(e^{j\omega}) \]

\[ \Phi_{ee}(e^{j\omega}) = \sigma_e^2/M \]
**Conclusion:** Thus, the quantization-noise power $\varepsilon \{x_{de}^2[n]\}$ has been reduced by a factor of $M$ through the filtering and downsampling, while the signal power has remained the same.

$$
\varepsilon \{x_{de}^2\} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{\sigma_e^2}{M} dw = \frac{\sigma_e^2}{M} = \frac{\Delta^2}{12M}
$$

For a given quantization noise power, there is a clear tradeoff between the oversampling factor $M$ and the quantization step $\Delta$. 

**Noise power reduction**
Oversampling for noise power reduction

- Remember that \( \Delta = \frac{X^m}{2^B} \)

- Therefore
  \[
  \mathcal{E}\{x_{de}^2\} = \frac{1}{12M} \left( \frac{X^m}{2^B} \right)^2
  \]

- The above equation shows that for a fixed quantizer, the noise power can be decreased by increasing the oversampling ratio \( M \).

- Since the signal power is independent of \( M \), increasing \( M \) will increase the signal-to-quantization-noise ratio.
Tradeoff between oversampling and quantization bits

- Alternatively, for a fixed quantization noise power, the required value for $B$ is

$$P_{de} = \mathcal{E}\{x_{de}^2\} = \frac{1}{12M} \left( \frac{X_m}{2^B} \right)^2$$

- From the equation, every doubling of the oversampling ratio $M$, we need $\frac{1}{2}$ bit less to achieve a given signal-to-quantization-noise ratio.

- In other words, if we oversample by a factor $M=4$, we need one less bit to achieve a desired accuracy in representing the signal.
Previously, we have shown that oversampling and decimation can improve the signal-to-quantization-noise ratio.

The result is remarkable, but if we want to make a significant reduction, we need very large sampling ratios.
- Eg., to reduce the number of bits from 16 to 12 would require \( M=4^4=256 \).

The basic concept in noise shaping is to modify the A/D conversion procedure so that the power density spectrum of the quantization noise is no longer uniform.
The noise-shaping quantizer, generally referred to as a **sampled-data Delta-Sigma modulator**, is roughly shown as the following figures.

- Analog form

![Diagram of Delta-Sigma Modulator](image-url)
Oversampled Quantizer with Noise Shaping

- Can be represented by the discrete-time equivalent system as follows:
  - Discrete-time form

\[
\begin{align*}
H(z) &= \frac{1}{1 - z^{-1}} \\
\omega_c &= \pi / M \\
x_d[n] &= x_{dc}[n] + x_{de}[n]
\end{align*}
\]
Modeling the quantization error

- As before, we model the quantization error as an additive noise source.
- Hence, the above figure can be replaced by the following linear model:

\[ x_a(t) \rightarrow C/D \rightarrow x[n] \rightarrow \frac{1}{1 - z^{-1}} \rightarrow y[n] \rightarrow e[n] \rightarrow \text{LPF} \rightarrow \downarrow M \rightarrow x_d[n] \]
This linear system has two inputs, \( x[n] \) and \( y[n] \). According to the linearity, we can get the output \( y[n] \) by

1. set \( x[n] = 0 \), find the output \( y[n] \) w.r.t. \( e[n] \)
2. set \( e[n] = 0 \), find the output \( y[n] \) w.r.t. \( x[n] \)
3. add the above two outputs.
Consider the output in the z-domain. We denote the transfer function from $x[n]$ to $y[n]$ as $H_x(z)$ and from $e[n]$ to $y[n]$ as $H_e(z)$.

When $e[n]=0$:

$$\frac{X(z) - Z^{-1}Y(z)}{1 - z^{-1}}$$
Output when $e[n]=0$

- We have

$$Y[z] = \frac{X[z] - z^{-1}Y(z)}{1 - z^{-1}}$$

- So

$$Y[z] - z^{-1}Y[z] = X[z] - z^{-1}Y(z)$$

- That is

$$Y[z] = X[z]$$

when $E[z]$ is zero.
When $x[n] = 0$:
Output when $x[n] = 0$

We have

$$Y[z] = e(z) - \frac{z^{-1}Y(z)}{1 - z^{-1}}$$

So

$$Y[z] - z^{-1}Y[z] = E[z] - z^{-1}E[z] - z^{-1}Y(z)$$

That is

$$Y[z] = (1 - z^{-1})E[z]$$

when $X[z]$ is zero.
In fact, feedback systems have been widely used (serves as a fundamental architecture) in control engineering.

Generally:

![Feedback System Diagram]

Formula:

\[
\frac{Y(z)}{X(z)} = \frac{G(z)}{1 + G(z)H(z)}; \quad \frac{E(z)}{X(z)} = \frac{1}{1 + G(z)H(z)}
\]
Another way of derivation

From the feedback system formula, we can also obtain that

\[ H_x(z) = \frac{Y(z)}{X(z)} = \frac{G(z)}{1 + G(z)H(z)} = \frac{1}{1 - Z^{-1}} = 1 \]

\[ H_e(z) = \frac{Y(z)}{E(z)} = \frac{1}{1 + G(z)H(z)} = \frac{1}{Z^{-1}} = 1 - Z^{-1} \]
Hence, in the time domain, we have

\[ y_x[n] = x[n] \]

\[ y_e[n] = \hat{e}[n] = e[n] - e[n-1] \]

Therefore, the output \( y[n] \) can be represented equivalently as

\[ y[n] = y_x[n] + y_e[n] = x[n] + \hat{e}[n] \]

The quantization noise \( e[n] \) has been modified as \( \hat{e}[n] \)
Power spectral density of the modified noise

To show the reduction of the quantization noise, let's consider the power spectral density of $\hat{e}[n]$. Since we have the input-output relationship between $e[n]$ and $\hat{e}[n]$ as

$$Y_e(z) = (1 - Z^{-1})E(z)$$

In the frequency domain, we have

$$\hat{E}(e^{jw}) = Y_e(e^{jw}) = (1 - e^{-jw})E(e^{jw})$$
Equivalent system

\[ e[n] \]

\[ 1 - z^{-1} \]

\[ \hat{e}[n] \]

\[ y[n] = x[n] + \hat{e}[n] \]

\[ T = \frac{\pi}{\Omega_N M} \]

LPF

\[ \omega_c = \pi / M \]

\[ \downarrow M \]

\[ x_d[n] \]
Power spectral density of the modified noise

The power-spectral-density relation between the modified and original quantization noises is thus

$$
\phi_{\text{ee}}(e^{jw}) = \| 1 - e^{-jw} \|^2 \phi_{ee}(e^{jw}) = \| 1 - e^{-jw} \|^2 \sigma_e^2
$$

$$
= (1 - e^{-jw})(1 - e^{jw})\sigma_e^2 = (1 - e^{-jw} - e^{-jw} + 1)\sigma_e^2
$$

$$
= (2 - (e^{-jw} + e^{-jw}))\sigma_e^2 = (2 - 2\cos(w))\sigma_e^2
$$

$$
= (2\sin(w/2))^2 \sigma_e^2
$$

\text{Model the original quantization error as white noise with this variance}

\text{p.s.d. of the modified noise}

\text{p.s.d. of the original noise}
Quantization-noise power

Remember that the downsampler does not remove any of the signal power, the signal power in $x_{da}[n]$ is

$$P_{da} = \mathcal{E}\{x_{da}^2[n]\} = \mathcal{E}\{x^2[n]\} = \mathcal{E}\{x_a^2(t)\}$$

The quantization-noise power in the final output is

$$P_{de} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{x_de x_{de}}(e^{jw})$$

See the following illustration for its computation.
Before decimation

Power spectral density of modified noise

$$\Phi_{\hat{e}\hat{e}}(e^{j\omega}) = 4\sigma_e^2 \sin^2 (\omega/2)$$

The modified noise density is non-uniform and lower in the effective band region.

$$\frac{1}{T} = \frac{\Omega_NM}{\pi}$$

$$\Phi_{xx}(e^{j\omega})$$

$$\sigma_e^2$$
After decimation

\[ \frac{1}{T'} = \frac{\Omega_N}{\pi} \]

\[ \Phi_{x_{de}x_{de}}(e^{j\omega}) = \frac{4\sigma^2_e}{M} \sin^2 \left(\frac{\omega}{2M}\right) \]

Down-scaled by M and also stretched by M.
Quantization-noise power

- Hence, the quantization-noise power in the final output is

\[ P_{de} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_{x_{de}x_{de}}(e^{jw}) \]

\[ = \frac{1}{2\pi} \frac{\Delta^2}{12M} \int_{-\pi}^{\pi} (2\sin\left(\frac{2}{2M}\right))^2 dw \]

- Assume that M is sufficiently large, we can approximate that

\[ = \sin\left(\frac{w}{2M}\right) \approx \frac{w}{2M} \]
Bits and quantization tradeoff in noise shaping

- With this approximation,

\[ P_{de} = \frac{1}{36} \frac{\Delta^2 \pi^2}{M^3} \]

- For a \((B+1)\)-bit quantizer and maximum input signal level between plus and minus \(X_m\), \(\Delta = X_m/2^B\). To achieve a given quantization-noise power \(P_{de}\), we have

\[ B = -\frac{3}{2} \log_2 M + \frac{1}{2} \log_2 (\pi / 6) - \frac{1}{2} \log_2 P_{de} + \log_2 X_m \]

- We see that, whereas with direct quantization a doubling of the oversampling ratio \(M\) gained \(\frac{1}{2}\) bit in quantization, the use of noise shaping results in a gain of 1.5 bits.
### TABLE 4.1  EQUIVALENT SAVINGS IN QUANTIZER BITS RELATIVE TO $M = 1$ FOR DIRECT QUANTIZATION AND FIRST-ORDER NOISE SHAPING

<table>
<thead>
<tr>
<th>$M$</th>
<th>Direct quantization</th>
<th>Noise shaping</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>1</td>
<td>2.2</td>
</tr>
<tr>
<td>8</td>
<td>1.5</td>
<td>3.7</td>
</tr>
<tr>
<td>16</td>
<td>2</td>
<td>5.1</td>
</tr>
<tr>
<td>32</td>
<td>2.5</td>
<td>6.6</td>
</tr>
<tr>
<td>64</td>
<td>3</td>
<td>8.1</td>
</tr>
</tbody>
</table>
The noise-shaping strategy can be extended by incorporating a second stage of accumulation, as shown in the following:

**Figure 4.66** Oversampled quantizer with second-order noise shaping.
Second-order (i.e., 2-stage) noise shaping

- In the two-stage case, it can be derived that

\[ H_e(z) = (1 - z^{-1})^2 \]

\[ \phi_{\hat{e}\hat{e}}(e^{jw}) = \sigma_e^2 (2\sin(w/2))^4 \]

- In general, if we extend the case to p-stages, the corresponding noise shaping is given by

\[ \phi_{\hat{e}\hat{e}}(e^{jw}) = \sigma_e^2 (2\sin(w/2))^{2p} \]
By evaluation, with $p=2$ and $M=64$, we obtain almost 13 bits of increase in accuracy, suggesting that a 1-bit quantizer could achieve about 14-bit accuracy at the output of the decimator.

Although multiple feedback loops promise greatly increased quantization-noise reduction, they are not without problems. Specifically, for large values of $p$, there is an increased potential for instability and oscillations to occur.