Let a rv $Z$ have an MGF $g_Z(r)$ for $0 \leq r < r_+$ and mean $Z < 0$. By the Chernoff bound, for any $\alpha > 0$ and any $r \in (0, r_+)$,

$$\Pr\{Z \geq \alpha\} \leq g_Z(r) \exp(-r\alpha) = \exp(\gamma_Z(r) - r\alpha)$$

where $\gamma_Z(r) = \ln g_Z(r)$. If $Z$ is a sum $S_n = X_1 + \cdots + X_n$, of IID rv’s, then $\gamma_{S_n}(r) = n\gamma_X(r)$.

$$\Pr\{S_n \geq na\} \leq \min_r (\exp[n(\gamma_X(r) - ra)])$$

This is exponential in $n$ for fixed $a$ (i.e., $\gamma'(r) = a$). We are now interested in threshold crossings, i.e., $\Pr\{\bigcup_n (S_n \geq \alpha)\}$. As a preliminary step, we study how $\Pr\{S_n \geq \alpha\}$ varies with $n$ for fixed $\alpha$.

$$\Pr\{S_n \geq \alpha\} \leq \min_r (\exp[n\gamma_X(r) - ra])$$

Here the minimizing $r$ varies with $n$ (i.e., $\gamma'(r) = a/n$).
\[
\Pr\{S_n \geq \alpha\} \leq \min_{0 < r < r_+} \exp\left(-\alpha \left[ r - \frac{n}{\alpha} \gamma_X(r) \right]\right)
\]

When \( n \) is very large, the slope \( \frac{\alpha}{n} = \gamma'(r_0) \) is close to 0 and the horizontal intercept (the negative exponent) is very large. As \( n \) decreases, the intercept decreases to \( r^* \) and then increases again.

Thus \( \Pr\{\bigcup_{n} \{S_n \geq \alpha\}\} \approx \exp(-\alpha r^*) \), where the nature of the approximation will be explained in terms of the Wald identity.

**Wald’s identity with 2 thresholds**

Consider a random walk \( \{S_n; n \geq 1\} \) with \( S_n = X_1 + \cdots + X_n \) and assume that \( X \) is not identically zero and has a semi-invariant MGF \( \gamma(r) \) for \( r \in (r_-, r_+) \) with \( r_- < 0 < r_+ \). Let \( \alpha > 0 \) and \( \beta < 0 \) be two thresholds. Let \( J \) be the smallest \( n \) for which either \( S_n \geq \alpha \) or \( S_n \leq \beta \).

Note that \( J \) is a stopping trial, i.e., \( I_{J=n} \) is a function of \( S_1, \ldots, S_n \) and \( J \) is a rv. The fact that \( J \) is a rv is proved in Lemma 7.5.1, but is almost obvious.

Wald’s identity now says that for any \( r, r_- < r < r_+ \),

\[
E[\exp(r S_J - J \gamma(r))] = 1.
\]

If we replace \( J \) by a fixed step \( n \), this just says that \( E[\exp(r S_n)] = \exp(n \gamma(r)) \), so this is not totally implausible.
E[\exp(rS_J - J\gamma(r))] = 1 \quad (\text{Wald's identity}).

Before justifying this, we use it to bound the probability of crossing a threshold.

**Corollary:** Assume further that $X < 0$ and that $r^* > 0$ exists such that $\gamma(r^*) = 0$. Then

$$\Pr\{S_J \geq \alpha\} \leq \exp(-r^*\alpha).$$

**Wald's id. at $r^*$** is $E[\exp(r^*S_J)] = 1$. Since $\exp(r^*S_J) \geq 0$,

$$\Pr\{S_J \geq \alpha\} E[\exp(r^*S_J) \mid S_J \geq \alpha] \leq E[\exp(r^*S_J)] = 1.$$

For $S_J \geq \alpha$, we have $\exp(r^*S_J) \geq \exp(r^*\alpha)$. Thus

$$\Pr\{S_J \geq \alpha\} \exp(r^*\alpha) \leq 1.$$

This is valid for all choices of $\beta < 0$, so it turns out to be valid without a lower threshold, i.e., $\Pr\{\bigcup_n \{S_n \geq \alpha\}\} \leq \exp(-r^*\alpha)$.

We saw before that $\Pr\{S_n \geq \alpha\} \leq \exp(-\alpha r^*)$ for all $n$, but this corollary makes the stronger and cleaner statement that $\Pr\{\bigcup_{n \geq 1} \{S_n \geq \alpha\}\} \leq \exp(-r^*\alpha)$.

The Chernoff bound has the advantage of showing that the $n$ for which the probability of threshold crossing is essentially highest is $n = \alpha/\gamma'(r^*)$. 
The Kingman bound for G/G/1

The corollary can be applied to the queueing time $W_i$ for the $i$th arrival to a G/G/1 system.

We let $U_i = X_i - Y_{i-1}$, i.e., $U_i$ is the difference between the $i$th interarrival time and the previous service time.

Recall that we showed that $\{U_i; i \geq 1\}$ is a modification of a random walk. The text shows that it is a random walk looking backward.

Letting $\gamma(r)$ be the semi-invariant MGF of each $U_i$, then the Kingman bound (the corollary to the Wald identity for the G/G/1 queue) says that for all $n \geq 1$,

$$\Pr\{W_n \geq \alpha\} \leq \Pr\{W \geq \alpha\} \leq \exp(-r^*\alpha); \quad \text{for all } \alpha > 0.$$ 

Large deviations for hypothesis tests

Let $\tilde{Y} = (Y_1, \ldots, Y_n)$ be IID conditional on $H_0$ and also IID conditional on $H_1$. Then

$$\ln(\Lambda(\tilde{y})) = \ln \frac{f(\tilde{y} \mid H_0)}{f(\tilde{y} \mid H_1)} = \sum_{i=1}^{n} \ln \frac{f(y_i \mid H_0)}{f(y_i \mid H_1)}$$

Define $z_i$ by $z_i = \ln \frac{f(y_i \mid H_0)}{f(y_i \mid H_1)}$

A threshold test compares $\sum_{i=1}^{n} z_i$ with $\ln(\eta) = \ln(p_1/p_0)$.

Conditional on $H_1$, make error if $\sum_i Z_i^1 > \ln(\eta)$ where $Z_i^1$, $1 \leq i \leq n$, are IID conditional on $H_1$. 

Exponential bound for $\sum_i Z_i^1$

$$\gamma_1(r) = \ln \left\{ \int f(y \mid H_1) \exp \left[ r \ln \frac{f(y \mid H_0)}{f(y \mid H_1)} \right] dy \right\}$$

$$= \ln \left\{ \int f^{1-r}(y \mid H_1) f^r(y \mid H_0) dy \right\}$$

At $r = 1$, this is $\ln(\int f(y \mid H_0) dy) = 0$. 

\[ q_1(\eta) \leq \exp n [\gamma_1(r_0) - r_0 \ln(\eta)/n] \]

where $q_{\ell}(\eta) = \Pr\{e \mid H = \ell\}$

---

Exponential bound for $\sum_i Z_i^0$

$$\gamma_0(s) = \ln \left\{ \int f(y \mid 0) \exp \left[ s \ln \frac{f(y \mid H_0)}{f(y \mid H_1)} \right] dy \right\}$$

$$= \ln \left\{ \int f^{-s}(y \mid H_1) f^{1+s}(y \mid H_0) dy \right\}$$

At $s = -1$, this is $\ln(\int f(y \mid H_1) dy) = 0$. Note: $\gamma_0(s) = \gamma_1(r-1)$.

\[ q_0(\eta) \leq \exp n [\gamma_1(r_0) + (1-r_0) \ln(\eta)/n] \]
These are the exponents for the two kinds of errors. This can be viewed as a large deviation form of Neyman Pearson. Choose one exponent and the other is given by the inverted see-saw above.

The a priori probabilities are usually not the essential characteristic here, but the bound for MAP is optimized at \( r \) such that \( \ln(\eta)/n - \gamma_0(r) \)

\[\gamma_0(r) - r \ln(\eta)/n\]
\[\gamma_1(r) + (1-r) \ln(\eta)/n\]

\[\ln[\Pr\{e | H_1\} / \Pr\{e | H_0\}] = \alpha\]

Also, \( \Pr\{e | H_1\} \leq e^\beta \).

**Sequential detection**

This large-deviation hypothesis-testing problem screams out for a variable number of trials.

We have two coupled random walks, one based on \( H_0 \) and one on \( H_1 \).

We use two thresholds, \( \alpha > 0 \) and \( \beta < 0 \). Note that \( E[Z | H_0] < 0 \) and \( E[Z | H_1] > 0 \).

Thus crossing \( \alpha \) is a rare event given the random walk with \( H_0 \) and crossing \( \beta \) is rare given \( H_1 \).

Since \( r^* = 1 \) for the \( H_0 \) walk, \( \Pr\{e | H_0\} \leq e^{-\alpha} \).

This is not surprising; for the simple RW with \( p_1 = 1/2 \), \( \sum_i Z_i = \alpha \) means that
Tilted probabilities

Let \( \{X_n; n \geq 1\} \) be a sequence of IID discrete random variables with a MGF at some given \( r \). Given the PMF of \( X \), define a tilted PMF (for \( X \)) as

\[
q_{X,r}(x) = p_X(x) \exp[r x - \gamma(r)].
\]

Summing over \( x \), \( \sum q_{X,r}(x) = g_X(r)e^{-\gamma_X(r)} = 1 \). We view \( q_{X,r}(x) \) as the PMF on \( X \) in a new probability space with this given relationship to the old space.

We can then use all the laws of probability in this new measure. In this new measure, \( \{X_n; n \geq 1\} \) are taken to be IID. The mean of \( X \) in this new space is

\[
E_r[X] = \sum_x x q_{X,r}(x) = \sum_x x p_X(x) \exp[r x - \gamma(r)]
\]

\[
= \frac{1}{g_X(r)} \sum_x \frac{d}{dr} p_X(x) \exp[r x]
\]

\[
= \frac{g'_X(r)}{g_X(r)} = \gamma'(r).
\]

The coupling between errors given \( H_1 \) and errors given \( H_0 \) is weaker here than for fixed \( n \).

Increasing \( \alpha \) lowers \( \Pr\{e | H_0\} \) exponentially and increases \( E[J | H_1] \approx \alpha / E[Z | H_1] \) (from Wald’s equality since \( \alpha \approx E[S_J | H = 1] \)). Thus

\[
\Pr\{e | H=0\} \sim \exp(-E[J | H=1] E[Z | H=1])
\]

In other words, \( \Pr \{e | H=0\} \) is essentially exponential in the expected number of trials given \( H=1 \). The exponent is \( E[Z | H=1] \), illustrated below.

Similarly, \( \Pr\{e | H=1\} \sim \exp(E[J | H=0] E[Z | H=0]) \).

Tilted probabilities

Let \( \{X_n; n \geq 1\} \) be a sequence of IID discrete random variables with a MGF at some given \( r \). Given the PMF of \( X \), define a tilted PMF (for \( X \)) as

\[
q_{X,r}(x) = p_X(x) \exp[r x - \gamma(r)].
\]

Summing over \( x \), \( \sum q_{X,r}(x) = g_X(r)e^{-\gamma_X(r)} = 1 \). We view \( q_{X,r}(x) \) as the PMF on \( X \) in a new probability space with this given relationship to the old space.

We can then use all the laws of probability in this new measure. In this new measure, \( \{X_n; n \geq 1\} \) are taken to be IID. The mean of \( X \) in this new space is

\[
E_r[X] = \sum_x x q_{X,r}(x) = \sum_x x p_X(x) \exp[r x - \gamma(r)]
\]

\[
= \frac{1}{g_X(r)} \sum_x \frac{d}{dr} p_X(x) \exp[r x]
\]

\[
= \frac{g'_X(r)}{g_X(r)} = \gamma'(r).
\]
The joint tilted PMF for $\tilde{X}^n = (X_1, \ldots, X_n)$ is then

$$q_{\tilde{X}^n,r}(x_1, \ldots, x_n) = \tilde{p}_{\tilde{X}^n}(x_1, \ldots, x_n) \exp(\sum_{i=1}^n [r x_i - \gamma(r)]).$$

Let $A(s_n)$ be the set of $n$-tuples such that $x_1 + \cdots + x_n = s_n$. Then (in the original space) $p_{S_n}(s_n) = \Pr\{S_n = s_n\} = \Pr\{A(s_n)\}$.

Also, for each $\tilde{x}^n \in A(s_n)$,

$$q_{\tilde{X}^n}(x_1, \ldots, x_n) = \tilde{p}_{\tilde{X}^n}(x_1, \ldots, x_n) \exp(rs_n - n\gamma(r))$$

$$q_{S_n,r}(s_n) = p_{S_n}(s_n) \exp[rs_n - n\gamma(r)],$$

where we have summed over $A(s_n)$. This is the key to much of large deviation theory. For $r > 0$, it tilts the probability measure on $S_n$ toward large values, and the laws of large numbers can be used on this tilted measure.

---

Proof of Wald’s identity

The stopping time $J$ for the 2 threshold RW is a rv (from Lemma 7.5.1) and it is also a rv for the tilted probability measure. Let $T_n = \{\tilde{x}_n : s_n \notin (\beta, \alpha); s_i \in (\beta, \alpha); 1 \leq i < n\}$.

That is, $T_n$ is the set of $n$ tuples for which stopping occurs on trial $n$. Letting $q_{J,r}(n)$ be the PMF of $J$ in the tilted probability measure,

$$q_{J,r}(n) = \sum_{\tilde{x}_n \in T_n} q_{\tilde{X}^n,r}(\tilde{x}_n) = \sum_{\tilde{x}_n \in T_n} \tilde{p}_{\tilde{X}^n}(\tilde{x}_n) \exp[rs_n - n\gamma(r)]$$

$$= E[\exp[rs_n - n\gamma(r) \mid J=n] \Pr\{J=n\}.$$ 

Summing over $n$ completes the proof.