GREENWOOD'S FORMULA, ETC.

In the j^{th} (narrow) interval let there be d_j failures among y_j there at risk. Then we have that $d_j \mid y_j \sim bin(y_j, p_j)$, where p_j is the probability estimated by d_j / y_j .

The KM estimator is $\hat{S}(t) = \prod_{j:t_j<t} \{1 - d_j / y_j\}$, so $\log \hat{S}(t) = \sum_{j:t_j<t} \log\{1 - d_j / y_j\}$, and the quantities $1 - d_j / y_j$ turn out to be approximately independent (conditionally or not — this part is tricky). Thus $\operatorname{var} \hat{S}(t) \doteq \sum_{j:t_j<t} \operatorname{var}\{\log(1 - d_j / y_j)\}$. Using the "delta" method we have that $SD\{\log(1 - d_j / y_j)\} \doteq SD(1 - d_j / y_j) / q_j = \{p_j q_j / y_j\}^{1/2} / q_j$, so that $\operatorname{var}\{-\log \hat{S}(t)\} \doteq \sum_{j:t_j<t} \{p_j q_j / y_j\} / q_j^{\,2} = p_j / (y_j q_j)$. Replacing parameters by estimators yields that $\operatorname{var}\{-\log \hat{S}(t)\} \doteq \sum_{j:t_j<t} (d_j / y_j) / \{y_j (y_j - d_j) / y_j\} = \sum_{j:t_j<t} d_j / \{y_j (y_j - d_j)\}$.

If we used instead the Poisson approximation $d_j \mid y_j \sim po(y_j p_j)$, essentially the same argument yields the result $\operatorname{var}\{-\log \hat{S}(t)\} \doteq \sum_{j:t_j<t} d_j / y_j^{\,2}$. It is useful to note that this results from replacing the binomial variance $\operatorname{var}(d_j \mid y_j) = y_j p_j q_j$ by the Poisson variance $\operatorname{var}(d_j \mid y_j) = y_j p_j$.

But there is more to this alternative final expression. It follows from martingale theory that the quantities $d_j / y_j = dN_{\bullet}(s) / Y_{\bullet}(s)$ in the NA estimator, which are only approximately independent as used above, are exactly uncorrelated. With the NA estimator the delta method is not needed, and it follows further from the martingale theory that $\operatorname{var} \hat{\Lambda}(t) = E\left\{ \int_0^t \frac{1}{Y_{\bullet}(s)} \lambda(s) ds \right\}$ exactly. Note that within the expectation the $Y_{\bullet}(s)$ remains random, whereas we treated this as fixed in the above argument. This sort of "expectation equals expectation" expression is typical in martingale theory. Dropping the expectation operation and replacing $\lambda(s)$ by its NA estimator $dN_{\bullet}(t) / Y_{\bullet}(t)$ yields the approximation of the notes $\operatorname{var}\{\hat{\Lambda}_{NA}(t)\} \doteq \int_0^t \frac{1}{Y_{\bullet}(s)^2} dN_{\bullet}(s)$.