It is probably better to start with the full level of generality. So let $F(x,u)p(u)$ be the likelihood with prior where we ignore the observed data. Here u are the random effects and x are the fixed effects including possibly a parameteriozation for the covariance structure of the random effects. Let $g(x,u)$ be any smooth function of x and u. Given the observed data we want to make inferences on the value of g(x,u), such at the probability that $g(x,u) = t$. (We really mean here the probability that $g(x,u)$ is close to t.)

Given all the misinformation about this problem on the R list, it is in fact surprisingly easy to formulate a solution at least in principle. This solution leads to a practical solution obtained by several applications of the Laplace approximation. The confusion seems to arise from the fact that with empirical Bayes the random effects are integrated out so in a sense they no longer exist. However there is no reason why we need to be restricted to the standard empirical Bayes procedure if we are interested in the properties of a particular function of the fixed and random effects. We can in effect tailor the estimation procedure to the function of interest.

Consider the simple case where only one random effect is contained in the function of interest for example $g(x,u) = \sigma u_1$. We can integrate out the parameters $u_2, \ldots, u_n$ to obtain

$$R(x,u_1) = \frac{\int F(x,u)p(u)du_2du_3 \ldots du_n}{\int F(x,u)p(u)du_1du_2du_3 \ldots du_n}$$

For fixed x, $R(x,u_1)$ is a probability density function for u_1. Now fix a number t and consider the probability that $t - \epsilon/2 < \sigma u_1 < t + \epsilon/2$. This is just the probability that $(t-\epsilon/2)/\sigma < u_1 < (t+\epsilon/2)/\sigma$. which is equal to $\epsilon R(x,t/\sigma)/\sigma + O(\epsilon^2)$. Letting $\epsilon \to 0$ yields the probability density function $r(x,t) = R(x,t/\sigma)/\sigma$ for the function σu_1. To make inferences about σu_1 we profile over $r(x,t)$ to get the profile likelihood $P(t)$.

$$P(t) = \max_{\{x:\sigma u_1 = t\}} r(x,t)$$

.

The general case is almost identical except that the integration is subtler. For any fixed value of x we have a probability function for u, $q(x,u)$ give by

$$q(x,u) = \frac{F(x,u)p(u)}{\int F(x,u)p(u)du}$$

The denominator $\int F(x,u)p(u)du$ can be calculated via the standard Laplace approximation we already use. For any open set Ω the probability thant $u \in \Omega$ is

$$\int_\Omega q(x,u)du$$

We want to consider this integral for open sets of a special type determined by the function $g(x,u)$ for fixed x. Fix a value of t and let $\Omega_{\epsilon,t} = \{u : t - \epsilon/2 < g(x,u) < t + \epsilon/2\}$ Let

$$r(x,t) = \lim_{\epsilon \to 0} \frac{1}{\epsilon} \int_{\Omega_{\epsilon,t}} q(x,u)du$$

For a fixed value of x, $r(x,t)$ is a probability density function for t. It is r(x,t) which we will profile over t to make inferences about $g(x,u)$. That is the profile likelihood $P(t)$ for $g(x,u)=t$ will be

$$P(t) = \max_{\{x:g(x,u)=t\}} r(x,t)$$

Of course this assumes that we can calulate $r(x,t)$ which is impossible for most practical problems. Instead We shall use a form of the Laplace approximation. Let

$$\hat{u}(x,t) = \max_{\{u:g(x,u)=t\}} F(x,u)p(u)$$

We want to consider the part of the region in $\Omega_{\epsilon,t}$ near $\hat{u}(x,t)$ for small ϵ. In the direction $\nabla_u g(x,\hat{u}(x,t))$ it is very thin with width

$$\frac{\epsilon}{||\nabla_u g(x,\hat{u}(x,t))||} + O(\epsilon^2)$$

. Since $\nabla_u g(x,\hat{u}(x,t)) = \lambda \nabla_u F(x,\hat{u}(x,t))p(\hat{u}(x,t))$ for some nonzero number λ (in the nondegenerate case) it follow that if we restrict ourselves to directions perpendicular to $\nabla_u g(x,\hat{u}(x,t))$ then $F(x,u)$ has a critical point (maximum) and we can approximate it by its second order taylor expansion in these directions and integrate via the Laplace approximation. To do this we need the determinant. I don't know the best way to do this but one can do it as follows. Start with the standard basis e_i and form all the dot products $k = \nabla_u g(x,\hat{u}(x,t))/||\nabla_u g(x,\hat{u}(x,t))||$, and $f_i =< e_i - < e_i, k > k$. Discard the f_i with the smallest norm so that now there are $n-1$ of them.

$$S_{ij} = D^2_{uu} F(x,\hat{u}(x,t))p(\hat{u}(x,t))(f_i,f_j) \quad \text{for } 1 \le i \le n-1, 1 \le j \le n-1$$

and

$$H_{ij} =< f_i, f_j > \quad \text{for } 1 \le i \le n-1, 1 \le j \le n-1$$

The ratio of the determinants of these matrices

$$\frac{||S||}{||H||}$$

is what you want (or maybe the square root of it.)

Now how does this work for only one random effect. In that cas the level set of the function is only a point. Consider the function $\sigma_u u_1$. the norm of the gradient is σ, so the width is ϵ/σ_u and taking the derivative wrt ϵ we get $1/\sigma_u$. So the gradient of the function contains the information about the variance. Generalizing to n dimensions it would appear that the gradient will contain the extra information to augment that gained from the Laplace approximation in the $n-1$ dimensional space.