A New Approximation For The Risk Premium With Large Risks*

Abstract

It is well known that the Arrow-Pratt approximation to the risk premium is only valid for small risks. In this paper we consider a second approximation that works well for both large and small risks, and in particular appears to be significantly superior to Arrow-Pratt for approximating the true value of the risk premium when the risk is large. The new approximation is based on risk-neutral probabilities.

JEL Classification N^o: D5

Keywords: risk premium, Arrow-Pratt approximation, risk neutral probabilities

Richard Watt
Department of Economics, University of Canterbury, Private Bag 4800, Christchurch, NEW ZEALAND

*VERY PRELIMINARY FIRST DRAFT - PLEASE DO NOT QUOTE

1 Introduction

It is very well known that the Arrow-Pratt approximation to the true value of the risk premium (Pratt 1964, Arrow 1971) only works well whent the risk involved is "small". This of course is due to the fact that the Arrow-Pratt approximation depends upon the value of absolute risk aversion measured at the expected value of the risk in question, and absolute risk aversion is only a valid measure for risks that are close to certainty. Exactly how small the risk needs to be for the Arrow-Pratt approximation to be valid is unclear, but in general we can say that its value as an approximation deteriorates with the size of the risk (indeed, the deterioration is rapid, as we shall show in some simulations below). It appears to be an open question as to whether or not a better approximation can be found. In this paper we attempt such an exercise.

Before beginning, we do need to clarify what is meant by a "better" approximation. Naturally, if the utility function is known, then the true value of the risk premium can always be calculated exactly. The value of the Arrow-Pratt approximation is that it relies not on knowledge of the utility function per se, but rather of risk aversion. It is possible that risk aversion can be much more reliably guessed at than can an exact utility function, and therefore the Arrow-Pratt approximation is useful. A "better" approximation would need to satisfy at least two criteria – (1) it should provide a closer approximation to the true value of the risk premium, and (2) it should be based only indirectly on knowledge of the utility function itself, that is, it should be based on a measure that is related to the utility function, but for which we may have a reasonable ability to guess its value. The approximation suggested here appears to satisfy these two criteria, at least for large risks. The information that is required for the approximation to work is what is known in the financial literature as the "risk neutral probabilities", rather than risk aversion as in Arrow-Pratt.

2 The Basic Model

To see how the approximation that is suggested here comes about, we firstly consider some simple theory of optimal choice under risk. Assume an individual with the opportuntity to invest in a risk-free asset with payoff $\overline{w}$, and a risky asset denoted by $\widetilde{w}$. We assume explicitly that the expected value of the risky asset is strictly positive, $E\widetilde{w} > 0$ The individual may choose any convex combination of the two investment opportunities, and we denote the investment in the risky asset by λ. The investor's indirect utility function for money, $u(w)$, is assumed to be strictly increasing and concave. Concretely then, the investor chooses λ to maximise

$$U(\lambda) \equiv Eu((1-\lambda)\overline{w} + \lambda\widetilde{w})$$

The first order condition for an optimal investment is

$$U'(\lambda^*) = Eu'((1-\lambda^*)\overline{w} + \lambda^*\widetilde{w})(\widetilde{w} - \overline{w}) = 0$$

The second order condition is satisfied by the assumption of concavity.

Consider the two special cases of optima at extreme values of λ. Firstly, for the optimal investment to be $\lambda = 0$, we require (from the first order condition) that $Eu'(\overline{w})(\widetilde{w} - \overline{w}) = 0$, that is, $u'(\overline{w})(E\widetilde{w} - \overline{w}) = 0$. Thus, the investor will only choose to have a fully risk free final portfolio if the expected value of the risky investment is equal to the value of the risk-free endowment, something that we have assumed away. Now, consider when the investor would set $\lambda = 1$, that is, she would choose to have no risk free asset in the final portfolio. For this to occur, we require $Eu'(\widetilde{w})(\widetilde{w}-\overline{w}) = 0$. Since, by our assumption of concavity, $Eu'((1-\lambda)\overline{w}+\lambda\widetilde{w})(\widetilde{w}-\overline{w})$ is decreasing in λ, we get the following results:

1) $\lambda^* > 0$ if and only if $E\widetilde{w} > \overline{w}$, as has been assumed (and so $\lambda^* = 0$ if and only if $E\widetilde{w} \leq \overline{w}$).

2) $\lambda^* = 1$ if and only if $Eu'(\widetilde{w})(\widetilde{w} - \overline{w}) \geq 0$ (and so $\lambda^* < 1$ if and only if $Eu'(\widetilde{w})(\widetilde{w} - \overline{w}) < 0$).

Note that the condition for $\lambda^* = 1$ can more usefully be expressed as

$$\overline{w} \leq \frac{Eu'(\widetilde{w})\widetilde{w}}{Eu'(\widetilde{w})} \equiv \widehat{w}$$

Lemma 1 *$\widehat{w} < E\widetilde{w}$.*

In short, the investor's optimal strategy can be expressed in accordance with the exact value of risk free wealth that is involved in the choice. If $\overline{w} \leq \widehat{w}$, then $\lambda^* = 1$ (i.e., no sure wealth is retained in the optimal choice), if $\widehat{w} < \overline{w} < E\widetilde{w}$, then $0 < \lambda^* < 1$ (i.e. both risky and non-risky assets are retained), and if $\overline{w} \geq E\widetilde{w}$ then $\lambda^* = 0$ (i.e. no risk at all is retained in the optimal choice).

Readers may recognise that $\widehat{w}$ is the expected value of the risky asset valued at risk-neutral probabilities (see, for example, Gollier 2001, chapter 5). Besides Lemma 1, $\widehat{w}$ has some very desirable properties that are useful for measuring risk premia:

Lemma 2 *If $\widetilde{w}$ is degenerate, $\widetilde{w} = k$, then $\widehat{w} = k$.*

Lemma 3 *If the investor is risk neutral, then $\widehat{w} = E\widetilde{w}$.*

Lemma 4 *If $u_1(w)$ is more risk averse than $u_2(w)$, then $\widehat{w}_1 < \widehat{w}_2$, where $\widehat{w}_i$ is the value of $\widehat{w}$ corresponding to $u_i(w)$.*

Lemma 5 *A mean-preserving increase in the variance of $\widetilde{w}$ will decrease the value of $\widehat{w}$.*

Given Lemmas 1-5, it would seem that the value of $E\widetilde{w} - \widehat{w}$ (indeed, any positive transform of $E\widetilde{w} - \widehat{w}$) might be valuable as a measure of the risk premium corresponding to the risk $\widetilde{w}$. Under Lemmas 1 and 3, this risk premium must always be positive if and only if the investor is risk averse. Under Lemmas 1 and 2, this risk premium is valued at 0 when there is no risk, and positive otherwise (assuming risk aversion). Under Lemma 4, an increase in risk aversion will increase the value of this risk premium. And under Lemma 5, a pure increase in the underlying risk will also increase the value of this risk premium.

However, if we consider a two-dimensional setting, a suitable transformation of $E\widetilde{w} - \widehat{w}$ for the purposes of estimating the risk premium (as classically defined), becomes evident. If $\widetilde{w}$ is a two-dimensional risk, then it can be expressed as a payoff of, say, w_1^r with probability $1-p$, and a payoff of w_2^r with probability p (where the subindex r is to remind us that this is a risky asset) Thus we get:

$$\begin{aligned} \widehat{w} &= \frac{Eu'(\widetilde{w})\widetilde{w}}{Eu'(\widetilde{w})} \\ &= \frac{(1-p)u'(w_1^r)w_1^r + pu'(w_2^r)w_2^r}{(1-p)u'(w_1^r) + pu'(w_2^r)} \end{aligned}$$

On the other hand, the supporting tangent line to the indifference curve at the point $[w_1, w_2]$ is defined by:

$$w_2 - w_2^r = -\frac{(1-p)u'(w_1^r)}{pu'(w_2^r)}(w_1 - w_1^r)$$

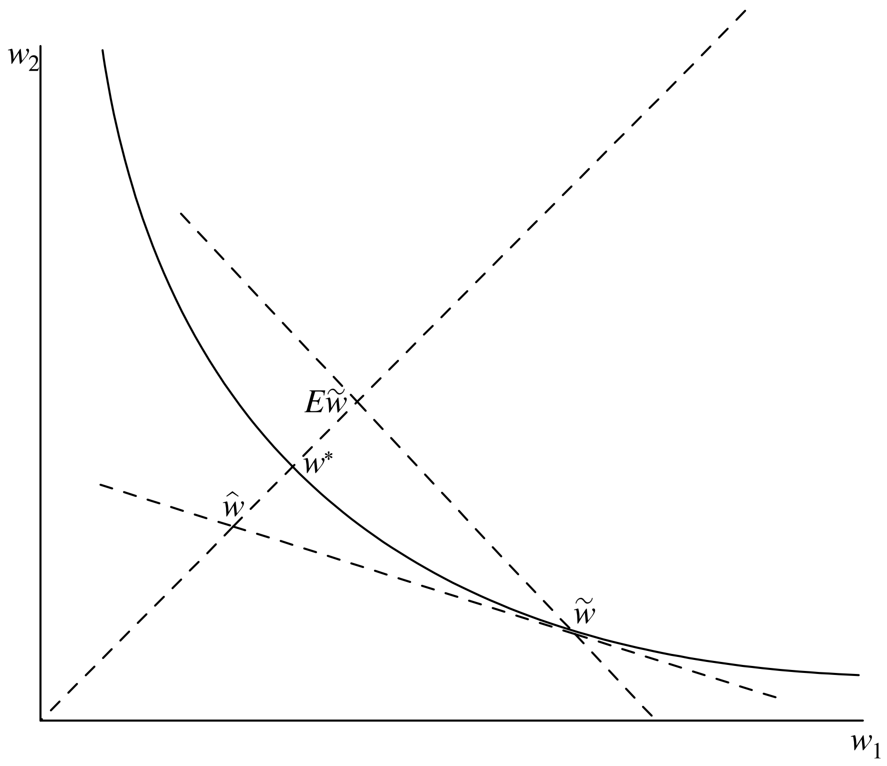

Figure 1: Figure 1: A two-dimensional example

Rearranging this, we get

$$(w_2 - w_2^r)pu'(w_2^r) = -(1-p)u'(w_1^r)(w_1 - w_1^r)$$

Setting $w_1 = w_2 \equiv w$, and collecting terms, this becomes

$$w((1-p)u'(w_1^r) + pu'(w_2^r)) = (1-p)u'(w_1^r)w_1^r + pu'(w_2^r)w_2^r$$

that is:

$$w = \frac{(1-p)u'(w_1^r)w_1^r + pu'(w_2^r)w_2^r}{(1-p)u'(w_1^r) + pu'(w_2^r)} = \widehat{w}$$

This is depicted in Figure 1.

Note from this graph that, in the case of a two-dimensional risk, $\widehat{w} < w^*$, where w^* is the certainty equivalent of the risk $\widetilde{w}$. This is, in fact, also a general theorem that is robust to any dimensionality as the next Lemma notes.

Lemma 6 $\widehat{w} < w^*$.

Of course, Lemma 1 is now really just a corollary of Lemma 6, since under strict risk aversion it is always true that $w^* < E\widetilde{w}$. In any case, Figure 1 clearly points to the estimate of the risk premium that we can use. The true risk premium corresponding to the risky situation $\widetilde{w}$ is defined as $\pi_t = E\widetilde{w} - w^*$. Without a doubt, the value $E\widetilde{w} - \widehat{w}$ is strictly greater than π_t, and so can be used as an upper bound for the true value of the risk premium, that is, $E\widetilde{w} - \widehat{w} > \pi_t > 0$. Thus, for any particular risk $\widetilde{w}$, and risk averse utility function $u(w)$, there exists some δ, $0 < \delta < 1$, such that $\pi_t = \delta(E\widetilde{w} - \widehat{w})$. As a first approximation, let us take $\delta = \frac{1}{2}$, which is equivalent to assuming that the certainty equivalent w^* lies exactly halfway between the expected value $E\widetilde{w}$ and the value $\widehat{w}$. Given that, our approximation to the true risk premium is:

$$\widehat{\pi} \equiv \frac{E\widetilde{w} - \widehat{w}}{2}$$

We now go on to compare this new approximation to the risk premium with the traditional Arrow-Pratt approximation, defined by $\pi_{ap} = \frac{\sigma_x}{2}\left(-\frac{u''(E\widetilde{w})}{u'(E\widetilde{w})}\right)$, where σ_x is the variance of the risky allocation. The easiest way to do this is by concrete examples.

3 Comparison of risk premium approximations

Let us retain the two-dimensional case for the time being, and to that assumption we add the assumption that $u(w) = Ln(w)$. Given this assumption on utility, we get

$$\begin{aligned} \widehat{w} &= \frac{Eu'(\widetilde{w})\widetilde{w}}{Eu'(\widetilde{w})} \\ &= \frac{E\left(\frac{1}{\widetilde{w}}\right)\widetilde{w}}{E\left(\frac{1}{\widetilde{w}}\right)} \\ &= \frac{1}{E\left(\frac{1}{\widetilde{w}}\right)} \end{aligned}$$

And so the assumption of two-dimensionality now gives us

$$\begin{aligned} \widehat{w} &\equiv \frac{1}{\frac{1-p}{w_1} + \frac{p}{w_2}} \\ &= \frac{1}{\frac{(1-p)w_2 + pw_1}{w_1 w_2}} \\ &= \frac{w_1 w_2}{(1-p)w_2 + pw_1} \end{aligned}$$

Finally, then, the new approximation to the risk premium for these assumptions is given by

$$\begin{aligned}\widehat{\pi} &\equiv \frac{E\widetilde{w}-\widehat{w}}{2} \\ &= \frac{1}{2}\left[(1-p)w_1 + pw_2 - \frac{w_1 w_2}{(1-p)w_2 + pw_1}\right]\end{aligned}$$

After some simplifiying steps, this reduces to

$$\widehat{\pi} = \frac{p(1-p)(w_1-w_2)^2}{2((1-p)w_2 + pw_1)}$$

On the other hand, since with log utility we get $-\frac{u''(E\widetilde{w})}{u'(E\widetilde{w})} = \frac{1}{E\widetilde{w}}$, and since with a two-dimensional risk the variance is $\sigma_x = p(1-p)(w_1 - w_2)^2$, it turns out that the Arrow-Pratt approximation for our assumptions is given by

$$\pi_{ap} = \frac{p(1-p)(w_1-w_2)^2}{2((1-p)w_1 + pw_2)}$$

which is very similar, but not identical to the other approximation.

Finally, for the particular case at hand, the true (i.e. the exact) value of the risk premium can also be calculated. It is, by definition, the value π_t that satisfies the equation $(1-p)Ln(w_1) + pLn(w_2) = Ln(E\widetilde{w} - \pi_t)$. This can be written as $Ln(w_1^{1-p}) + Ln(w_2^{p}) = Ln(E\widetilde{w} - \pi_t)$, that is $Ln(w_1^{1-p}w_2^{p}) = Ln(E\widetilde{w} - \pi_t)$. Therefore, the true risk premium is $\pi_t = E\widetilde{w} - w_1^{1-p}w_2^{p} = (1-p)w_1 + pw_2 - w_1^{1-p}w_2^{p}$.

In order to compare these three values, lets choose a "small" risk as $w_1 = 2, w_2 = 1$, and we graph the three functions as simple functions of the probability p. The three functions are:

$$\begin{aligned}\pi_t &= 2 - p - 2^{1-p} \\ \pi_{ap} &= \frac{p(1-p)}{2(2-p)} \\ \widehat{\pi} &= \frac{p(1-p)}{2(1+p)}\end{aligned}$$

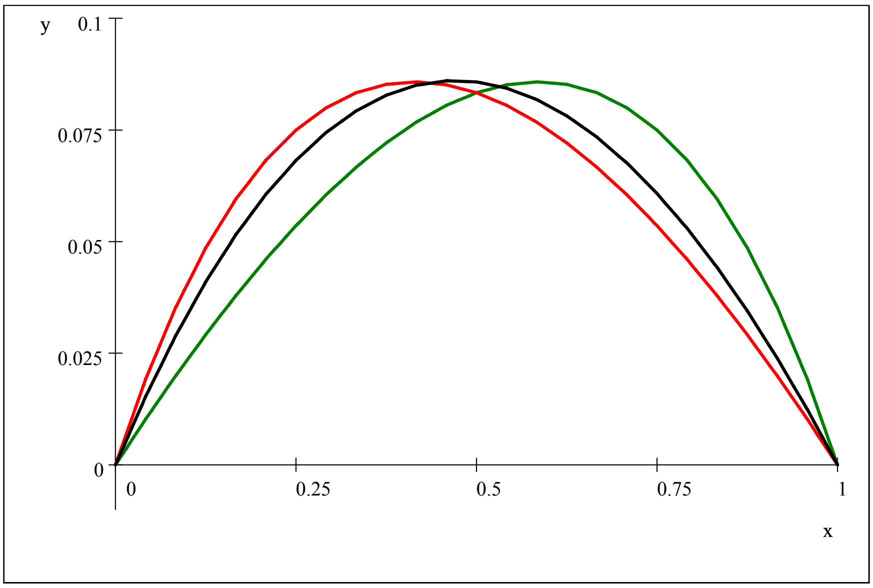

Figure 2: True risk premium (black curve), Arrow-Pratt approximation (green curve), and risk-neutral approximation (red curve), with binomial risk $x_1 = 2, x_2 = 1$ (x-axis represents the probability of state 2), assuming logarithmic utility.

This is done in Figure 2, where the black curve corresponds to π_t, the green one to π_{ap} and the red curve to $\widehat{\pi}$. Clearly, the approximation based on risk neutral probabilities tracks the true value of the risk premium somewhat better than does the Arrow-Pratt approximation, in spite of this being a relatively small risk.

To consider more exactly the improvement that the new approximation gives over the Arrow-Pratt one for this example, define the average error committed by the approximator $\pi(p, w)$, relative to the value to be approximated, $\pi_t(p, w)$, as

$$r(w_1, w_2)_\pi \equiv \int_0^1 \left(\frac{|\pi(p, w) - \pi_t(p, w)|}{\pi_t(p, w)} \right) dp$$

It turns out that $r(2, 1)_{\widehat{\pi}} = 0.11456$, while $r(2, 1)_{\pi_{ap}} = 0.23206$, that is, for the example at hand the Arrow-Pratt approximation is, on average, more than twice as bad as the new one for approximating the true risk premium.

Another reasonable measure of the relative success of the two measures is the proportion of values of probability p for which each is, respectively, the closest approximator. From Figure 2 it is clearly the case that for all values of p less than or equal to one half, the new approximator is better, and again for high values of p the new approximator is closer to the true value than is the Arrow-Pratt approximation. However, for a small range of p values just above one half, the Arrow-Pratt approximation does work better. However, even without calculating, the range of values for which Arrow-Pratt out-performs the new approximation is clearly smaller than the range over which the new one out-performs Arrow-Pratt.

Before leaving the two-dimensional case, consider the more general risk of $w_1 = z, w_2 = 1$, where $z \geq 2$. We get

$$\begin{aligned} \pi_t(z) &= (1-p)z + p - z^{1-p} \\ \widehat{\pi}(z) &= \frac{p(1-p)(z-1)^2}{2(1+p(z-1))} \\ \pi_{ap}(z) &= \frac{p(1-p)(z-1)^2}{2(z-p(z-1))} \end{aligned}$$

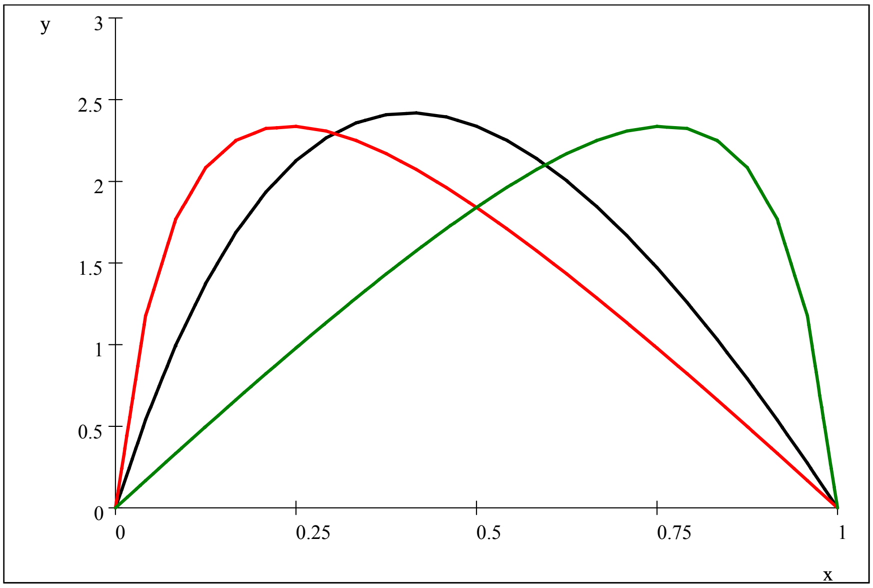

Figure 3: True risk premium (black curve), Arrow-Pratt approximation (green curve), and risk-neutral approximation (red curve), with binomial risk $x_1 = 10, x_2 = 1$ (x-axis represents the probability of state 2), assuming logarithmic utility.

These graphs are drawn in Figure 3 for the case of $z = 10$ (where again, the black curve is π_t, the red curve is $\widehat{\pi}$, and the green curve is π_{ap}). For the case of $z = 10$, we get $r(10,1)_{\widehat{\pi}} = 0.35201$, and $r(10,1)_{\pi_{ap}} = 0.80230$. For this example, although both approximations are worse in general than with the smaller risk, the Arrow-Pratt approximation is now significantly worse than two times less accurate than the new one. In table 1 we report the values of $r(z,1)$ for the two approximations, for a variety of values of z, as well as the value of $\frac{r(z,1)_{\widehat{\pi}}}{r(z,1)_{\pi_{ap}}}$.

Table 1: Goodness-of-fit			
z	$r(z,1)_{\widehat{\pi}}$	$r(z,1)_{\pi_{ap}}$	$\frac{r(z,1)_{\widehat{\pi}}}{r(z,1)_{\pi_{ap}}}$
2	0.11456	0.23206	0.49366
4	0.22358	0.47002	0.47568
6	0.28296	0.61409	0.46078
8	0.32268	0.71909	0.44873
10	0.35201	0.80230	0.43875
12	0.37498	0.87149	0.43028
120	0.59019	1.8419	0.32043

For all values of z the new approximation out-performs the Arrow-Pratt approximation, and the fact that $\frac{r(z,1)_{\widehat{\pi}}}{r(z,1)_{\pi_{ap}}}$ steadily declines as z increases indicates that as the risk gets bigger (in the sense used here), even though both approximations perform worse, the extent to which the performance of the new approximation is worsened is always less than that the extent to which the Arrow-Pratt approximation worsens.

As a second simulation, consider the case of a uniformly distributed risk. Specifically, let the risk be uniformly distributed between 1 and $z \geq 2$, so that the probability density function describing the risk is $f(z) = \frac{1}{z-1}$. Under this assumption, it is well known that the expected value of the risk is $E\widetilde{w} = \frac{z+1}{2}$, and that the variance is $\sigma^2(\widetilde{w}) = \frac{(z-1)^2}{12}$. Maintaining the assumption of logarithmic utility, the true value of the risk premium is given by the solution to

$$\int_1^z \frac{1}{z-1} Ln(w)dw = Ln(E\widetilde{w} - \pi_t)$$

After some simple calculations, this turns out to be

$$\begin{aligned}\pi_t(z) &= E\widetilde{w} - e^{\left(\frac{z}{z-1}Ln(z)-1\right)} \\ &= \frac{z+1}{2} - e^{\left(\frac{z}{z-1}Ln(z)-1\right)}\end{aligned}$$

Secondly, it is also easy to calculate the Arrow-Pratt approximation as

$$\begin{aligned}\pi_{ap}(z) &= \frac{\sigma^2(\widetilde{x})}{2}\left(\frac{1}{E\widetilde{w}}\right) \\ &= \frac{(z-1)^2}{12(z+1)}\end{aligned}$$

And finally, the new approximation under this scenario is

$$\begin{aligned}\widehat{\pi}(z) &\equiv \frac{E\widetilde{w} - \widehat{w}(z)}{2} \\ &= \frac{1}{2}\left(\frac{z+1}{2} - \frac{1}{E\left(\frac{1}{\widetilde{w}}\right)}\right)\end{aligned}$$

It can be easily checked that, under the uniform distribution assumption, we get $E\left(\frac{1}{\widetilde{w}}\right) = \frac{Ln(z)}{z-1}$, so that

$$\widehat{\pi}(z) = \frac{1}{2}\left(\frac{z+1}{2} - \frac{z-1}{Ln\,(z)}\right)$$

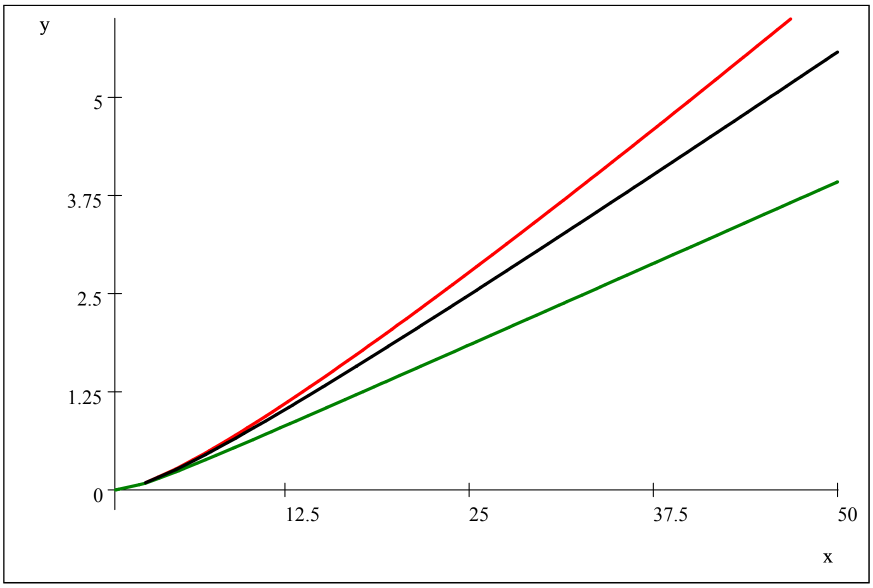

Figure 4: True risk premium (black curve), Arrow-Pratt approximation (green curve), and risk-neutral approximation (red curve), with uniformly distributed risk on $[1, z]$ (x-axis represents z), assuming logarithmic utility.

The three functions $\pi_t(z), \pi_{ap}(z)$ and $\widehat{\pi}(z)$ are drawn in Figure 4. As can be seen, the Arrow-Pratt approximation underestimates the true risk premium, while the new approximation over-estimates it. But it is also clear that, for all z, the new approximation does a somewhat better job at estimating the true risk-premium in terms of minimizing the absolute value of the error.

If we attempt these comparisons with a constant absolute risk aversion utility function, $u(w) = -e^{-Aw}$, (rather than the constant relative risk aversion one used above), we get the following results. We shall only do the exercise for the case of the uniform distribution. After some calculations, it is revealed that for this utility function:

$$\begin{aligned}
\pi_t &= \frac{\left(\left((z+1)Ae^{A}+2\right)e^{Az}-2e^{A}\right)e^{-A(z+1)}}{2A} = \frac{z+1}{2} + \frac{e^{-A}}{A} - e^{-z} \\
\pi_{ap} &= \frac{A(1-z)^2}{24} \\
\widehat{\pi} &= \frac{1}{2}\left(\frac{z+1}{2} - \left(\frac{\left(\frac{A+1}{A}\right)e^{-A} - \left(z+\frac{1}{A}\right)e^{-Az}}{e^{-A}-e^{-Az}}\right)\right)
\end{aligned}$$

The graphs of these three functions are drawn in Figure 5.

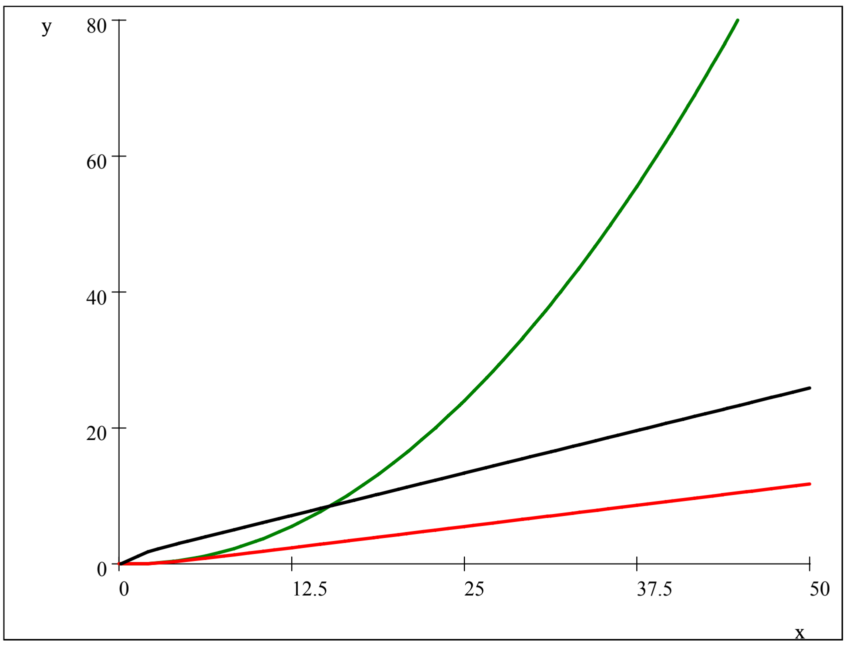

Figure 5: True risk premium (black curve), Arrow-Pratt approximation (green curve), and risk-neutral approximation (red curve), with uniformly distributed risk on $[1, z]$ (x-axis represents z), assuming CARA=1 utility.

As can be seen in Figure 5, the Arrow-Pratt approximator is the best for small levels of risk, but once the risk is over about 25, the Arrow-Pratt approximator spirals totally out of line with the true risk premium, while the new approximator is far better.the larger is the risk.

Of course, the above exercise can be repeated with any utility function and risk distribution combination imaginable, and it is impossible to check that for all such combinations the new approximation to the risk premium proposed here always works better than the Arrow-Pratt approximation. So no such claim is made, and it is left to the research agenda to attempt to search for a general theorem that compares the two approximations.

4 Conclusions

In this paper we have considered an alternative to the Arrow-Pratt approximation for estimating the value of the risk premium corresponding to a given risk. The new approximation is based on the expected value of the risk calculated at the "risk neutral" probabilities. We have compared this new approximation to the Arrow-Pratt one for a variety of situations, with different types of risk and different utility characteristics. The risk neutral approximation is better than the Arrow-Pratt approximation in all simulations when the risk is sufficiently large, and is also better for some simulations with small risk.

Whether or not the risk neutral approximation is as easy to calculate from real-world data than the Arrow-Pratt approximation is an empirical matter that needs to be looked into. But at least the information required to be able to calculate the new approximation is the answer to a single simple question – given the choice between any linear combination of the risk in question and an amount of sure wealth, what is the greatest amount of sure wealth such that the optimal choice is to have only the risk and no sure wealth in the final portfolio? As far as I know, there is no direct manner (e.g. a single question) to elicit the value of risk aversion as is required by the Arrow-Pratt approximation. So, at the very least, the new approximation appears to at least hold promise.

References

Arrow, K. (1971), Essays in the Theory of Risk Bearing, Chicago: Markham Publishing Co.

Gollier, C. (2001), The Economics of Risk and Time, Cambridge MA, MIT Press.

Pratt, J. (1964), "Risk Aversion in the Small and in the Large", *Econometrica*, 32; 122-36.