Start with the normalized SIR:

$$i = Bxy; \dot{x} = -i; \dot{y} = i - y$$

If $C(t)$ is the case reproductive number for an individual infected at time t, then the overall mean is:

$$\bar{C} = \frac{\int dt\, i(t)C(t)}{\int dt\, i(t)} = \frac{\int dt\, i(t)C(t)}{Z},$$

where Z is the size of the epidemic.

We can calculate $C(T)$ by integrating over time since infection δ:

$$C(t) = B \int d\delta\, x(t+\delta) \exp(-\delta),$$

or, in terms of time of contact $\tau = t + \delta$:

$$C(t) = B \int_{\tau > t} d\tau\, x(\tau) \exp(t - \tau),$$

We can also solve the $\dot{y}$ equation, by looking at who was infected at time θ, and how many of them survived:

$$y(t) = \int_{\theta < t} d\theta\, i(\theta) \exp(\theta - t),$$

and then expand:

$$\begin{aligned}
Z\bar{C} &= \int dt\, i(t)C(t) && (1)\\
&= B \int dt\, i(t) \int_{\tau > t} d\tau\, x(\tau) \exp(t - \tau) && (2)\\
&= B \int d\tau\, x(\tau) \int_{t < \tau} dt\, i(t) \exp(t - \tau) && (3)\\
&= B \int d\tau\, x(\tau) y(\tau) && (4)\\
&= \int d\tau\, i(\tau) && (5)\\
&= Z && (6)
\end{aligned}$$

Now try to expand the sum of squares. One trick we're using here is that to integrate $f^2(x)$ over a range, we can just take twice the integral of

$f(x)f(y)$ over a triangle where we assume we know $x > y$. This allows a similar limit-switching move as above.

Expanding was a disaster (not currently shown), so we try a half-expansion (expand one of the two factors of C).

$$
\begin{aligned}
& \int dt\, i(t)C^2(t) && (7)\\
= \quad & 2B \int dt\, i(t) \int_{\tau>t} d\tau\, x(\tau)\exp(t-\tau)C(\tau) && (8)\\
= \quad & 2B \int d\tau\, x(\tau)C(\tau) \int_{t<\tau} dt\, i(t)\exp(t-\tau) && (9)\\
= \quad & 2B \int d\tau\, x(\tau)C(\tau)y(\tau) && (10)\\
= \quad & 2 \int d\tau\, C(\tau)i(\tau) && (11)
\end{aligned}
$$

Wait! Does this prove the Roswell Conjecture‽ It *seems* too simple (we haven't seen how far the "proof" would go with a different infectiousness kernel), but maybe we accidentally leveraged the distribution-matching property of the exponential (infectious period has same distribution as generation interval) – that actually seems plausible.

PS: This does balance correctly, fwiw: it yields a squared-sum of $2Z/Z = 2$ to go with the mean of 1, so a variance of 1.