Median $\tilde{\mu}$: Solution to $F(x) = \frac{1}{2}$ (always exists)

gamma(k, β) distribution (time of k^{th} arrival)
$\mu = k\beta$, $\sigma = \sqrt{k}\beta$, $M(t) = (1 - \beta t)^{-k}$

$$f(x) = \frac{x^{k-1} \exp\left(-\frac{x}{\beta}\right)}{(k-1)!\beta^k} \qquad x > 0$$

Central Limit Theorem
RIS of size n from (*almost*) *any* distribution.

First, we standardize $\bar{X}$

$$Z = \frac{\bar{X} - \mu}{\frac{\sigma}{\sqrt{n}}} = \sum_{i=1}^{n} \frac{X_i - \mu}{\sigma\sqrt{n}} \equiv \sum_{i=1}^{n} U_i$$

The first 4 simple moments of each U_i are:

$$\begin{aligned}
\mathbb{E}(U) &= 0 \\
\mathbb{E}\left(U^2\right) &= \frac{1}{n} \\
\mathbb{E}\left(U^3\right) &= \frac{\mathbb{E}\left[(X-\mu)^3\right]}{\sigma^3 n^{3/2}} \\
\mathbb{E}\left(U^4\right) &= \frac{\mathbb{E}\left[(X-\mu)^4\right]}{\sigma^4 n^2}
\end{aligned}$$

This means that the corresponding MGF of Z is

$$M_Z(t) = \left(1 + \frac{1}{n} \cdot \frac{t^2}{2} + \frac{\text{SK}}{n^{3/2}} \cdot \frac{t^3}{3!} + \frac{\text{KT}}{n^2} \cdot \frac{t^4}{4!} + ...\right)^n$$

Our

$$\left(1 + \frac{a}{n} + \frac{b}{n^2} + ...\right)^n \to e^a$$

formula tells us that the MGF of Z has the following $n \to \infty$ limit

$$M_Z(t) \to \exp\left(\frac{t^2}{2}\right)$$

The corresponding pdf is

$$f(z) = \frac{\exp\left(-\frac{z^2}{2}\right)}{\sqrt{2\pi}} \qquad \text{any real } z$$

(yet to be proven).

First we verify that $\sqrt{2\pi}$ is the proper normalizing constant, by evaluating

$$\iint_{\text{whole plane}} \exp\left(-\frac{z_1^2+z_2^2}{2}\right) dz_1 dz_2$$

Introducing **polar coordinates** (a routine change of variables, whose **Jacobian** equals to r - a *must* to understand). In the new coordinates, the same integral becomes

$$\left(\int_0^{2\pi} d\theta\right) \cdot \left(\int_0^{\infty} \exp\left(-\tfrac{r^2}{2}\right) r\ dr\right)$$

Z is so called **standardized** Normal, or $\mathcal{N}(0,1)$.

'**General**' Normal RV

$$V = \sigma\ Z + \mu$$

Clearly, $\mathbb{E}(V) = \mu$, $\mathrm{Var}(V) = \sigma^2$. Notation: $\mathcal{N}(\mu, \sigma)$.

To get the pdf of V, we substitute

$$z = \frac{v-\mu}{\sigma}$$

in $f(z)$, and further divide by σ (since dz becomes $\dfrac{dv}{\sigma}$), getting

$$f(v) = \frac{\exp\left(-\dfrac{(v-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\cdot\sigma} \qquad (dv)$$

Similarly, the MGF of V is

$$M_V(t) = e^{\mu t} \cdot M_Z(t \to \sigma t) = \exp\left(\frac{\sigma^2 t^2}{2} + \mu t\right)$$

This implies that, adding two *independent*, Normally distributed RVs from $\mathcal{N}(\mu_1, \sigma_1)$ and $\mathcal{N}(\mu_2, \sigma_2)$, the sum is also Normally distributed with the mean equal to $\mu_1 + \mu_2$ and the standard deviation of $\sqrt{\sigma_1^2 + \sigma_2^2}$.

The Central Limit Theorem can now be stated as follows

$$\bar{X} \approx \mathcal{N}(\mu, \tfrac{\sigma}{\sqrt{n}})$$

or

$$\sum_{i=1}^{n} X_i \approx \mathcal{N}(n\mu, \sigma\sqrt{n})$$

so, to answer any probability question about these, we need to build the corresponding pdf and integrate (no tables!).

When the distribution we sample is discrete, to compute the probability of winning (in total) at least \$13 can be answered by either

$$\Pr\left(\sum_{i=1}^{n} W_i \geq 13\right) \qquad \text{or} \qquad \Pr\left(\sum_{i=1}^{n} W_i > 12\right)$$

Which value do we use in our Normal approximation?

Introducing **bivariate Normal** distribution (first step to multivariate)
First, we need the concept of **joint** MGF of two RVs, say X and Y

$$M(t_1, t_2) \equiv \mathbb{E}\left(e^{t_1 X \,+\, t_2 Y}\right)$$

What should be immediately obvious is that $M(t_1, t_2 = 0)$ yields (rather easily) the (**marginal**) MGF of X. Also, differentiating $M(t_1, t_2)$ w/r to t_1 yields

$$\mathbb{E}\left(X \cdot e^{t_1 X \,+\, t_2 Y}\right)$$

So, to get $\mathbb{E}(X)$, we have to set $t_1 = t_2 = 0$. In general

$$\mathbb{E}\left(X^2 Y\right) = \left.\frac{\partial^3 M(t_1, t_2)}{(\partial t_1)^2 \partial t_2}\right|_{t_t = t_2 = 0}$$

This means that, when Taylor-expanding $M(t_1, t_2)$, one gets

$$1 + \mu_x t_1 + \mu_y t_2 + \mathbb{E}\left(X^2\right) \frac{t_1^2}{2} + \mathbb{E}\left(Y^2\right) \frac{t_2^2}{2} + \mathbb{E}(X \cdot Y)\, t_1 t_2 +$$
$$\mathbb{E}\left(X^3\right) \frac{t_1^3}{3!} + \mathbb{E}\left(X^2 Y\right) \frac{t_1^2 t_2}{2!1!} + ...$$

Bivariate version of **CLT**
Consider a RIS of size n from a bivariate distribution (X is # of spades and Y # of diamonds when dealing 7 cards). We already know that both

$$Z_1 = \frac{\bar{X} - \mu_x}{\frac{\sigma_x}{\sqrt{n}}} \qquad \text{and} \qquad Z_2 = \frac{\bar{Y} - \mu_y}{\frac{\sigma_y}{\sqrt{n}}}$$

have, to a good approximation, $\mathcal{N}(0, 1)$ distribution, when n is large. The question is: what is their **joint** (bivariate) distribution?

And again, this can be settled only at the level of the corresponding *joint* MGF - let's try to build it:

$$\left(1 + \frac{1}{n} \cdot \frac{t_1^2}{2} + \frac{1}{n} \cdot \frac{t_2^2}{2} + \frac{\rho}{n} \cdot t_1 t_2 + \frac{...}{n^{3/2}} + ...\right)^n$$

whose $n \to \infty$ limit is

$$\exp\left(\frac{t_1^2 + t_2^2 + 2\rho\; t_1 t_2}{2}\right)$$

And, as before, what now is the corresponding bivariate pdf? Answer:

$$f(z_1, z_2) = \frac{\exp\left(-\dfrac{z_1^2 + z_2^2 - 2\rho\ z_1 z_2}{2(1-\rho^2)}\right)}{2\pi\sqrt{1-\rho^2}} \qquad \text{all plane}$$

Let's verify having the correct normalizing constant:

$$\int_{-\infty}^{\infty} \exp\left(-\frac{z_1^2}{2(1-\rho^2)}\right)\left[\int_{-\infty}^{\infty} \exp\left(-\frac{z_2^2 - 2\rho z_1 z_2}{2(1-\rho^2)}\right) dz_2\right] dz_1$$

By subtracting and adding $\rho^2 z_1^2$ to the 1st and 2nd denominator (respectively), we get $\sqrt{2\pi(1-\rho^2)}$ for the dz_2 integral and $\sqrt{2\pi}$ for the dz_1 integral (check).

We already know the two marginals (easy), how about a conditional pdf, say of Z_2 given that Z_1 has been observed to have the value of $\mathbf{z}_1$.

$$f(z_2|Z_1 = \mathbf{z}_1) = \frac{\exp\left(-\dfrac{\mathbf{z}_1^2 + z_2^2 - 2\rho\ \mathbf{z}_1 z_2}{2(1-\rho^2)}\right)}{2\pi\sqrt{1-\rho^2}} \div \frac{\exp\left(-\dfrac{\mathbf{z}_1^2}{2}\right)}{\sqrt{2\pi}}$$

$$= \frac{\exp\left(-\dfrac{z_2^2 - 2\rho\ \mathbf{z}_1 z_2 + \rho^2 \mathbf{z}_1^2}{2(1-\rho^2)}\right)}{\sqrt{2\pi}\sqrt{1-\rho^2}} = \frac{\exp\left(-\dfrac{(z_2 - \rho\ \mathbf{z}_1)^2}{2(1-\rho^2)}\right)}{\sqrt{2\pi}\sqrt{1-\rho^2}}$$

All we need to do now is to *identify* the answer as $\mathcal{N}(\rho\ \mathbf{z}_1, \sqrt{1-\rho^2})$.

To define '**general**' bivariate Normal, we linearly transform each Z_1 and Z_2, thus

$$\begin{aligned} U &= \sigma_1 Z_1 + \mu_1 \\ V &= \sigma_2 Z_2 + \mu_2 \end{aligned}$$

denoting the new distribution $\mathcal{N}(\mu_1, \mu_2, \sigma_1, \sigma_2, \rho)$. Note that the correlation coefficient between U and V stays the same ρ !

We know what the two (of U and V) marginals are, how about the conditional distribution of V given that $U = \mathbf{u}$? Well, this is the conditional distribution of $\sigma_2 Z_2 + \mu_2$ given that $\sigma_1 Z_1 + \mu_1 = \mathbf{u}$. We already know that the conditional distribution of Z_2 given that $Z_1 = \dfrac{\mathbf{u} - \mu_1}{\sigma_1}$ is

$$\mathcal{N}\left(\rho \cdot \frac{\mathbf{u} - \mu_1}{\sigma_1}, \sqrt{1-\rho^2}\right)$$

so, the conditional distribution of $\sigma_2 Z_2 + \mu_2$ is then

$$\mathcal{N}\left(\mu_2 + \sigma_2 \rho \cdot \frac{\mathbf{u} - \mu_1}{\sigma_1}, \sigma_2\sqrt{1-\rho^2}\right)$$