KNOW YOUR DISTRIBUTIONS!!

Discrete (integer-valued)

Univariate

Name	pmf	counts the number of
$\mathcal{B}(n,p)$	$\binom{n}{i}p^i q^{n-i}$	successes in n trials
$\mathcal{G}(p)$	pq^{i-1}	trials to get the first S
$\mathcal{NB}(k,p)$	$\binom{i-1}{k-1}p^k q^{i-k}$	trials to get k of them
$\mathcal{P}(\Lambda)$	$\frac{\Lambda^i}{i!}\cdot e^{-\Lambda}$	customers in fixed time
$\mathcal{HG}(n,K,N)$	$\frac{\binom{K}{i}\cdot\binom{N-K}{n-i}}{\binom{N}{n}}$	spades in a hand of n

formulas for mean, variance, PGF (support, modified $\mathcal{G}$ and $\mathcal{NB}$).
Important result: $P(z)^n$ is the PGF of an independent sum of n such RVs (no need for CLT)!

Multivariate

Multi-	pmf	Covariance
nomial	$\binom{n}{i,j,k}p_1^i p_2^j p_3^k$	$-np_1p_2$
variate$\mathcal{HG}$	$\frac{\binom{K_1}{i}\cdot\binom{K_2}{j}\cdot\binom{K_3}{k}}{\binom{N}{n}}$	$-n\cdot\frac{K_1K_2}{N^2}\cdot\frac{N-n}{N-1}$

Continuous (real-valued)

Univariate

Name	pdf	relates to
$\mathcal{U}(a,b)$	$\frac{1}{b-a}$	spinning wheel
$\mathcal{E}(\beta)$	$\frac{\exp\left(-\frac{x}{\beta}\right)}{\beta}$	time of next arrival
gamma(α,β)	$\frac{x^{\alpha-1}\exp\left(-\frac{x}{\beta}\right)}{\Gamma(\alpha)\beta^\alpha}$	time of α^{th} arrival
$\mathcal{N}(\mu,\sigma)$	$\frac{\exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)}{\sqrt{2\pi}\cdot\sigma}$	CLT
beta(k,m)	$\frac{\Gamma(k+m)\cdot x^{k-1}(1-x)^{m-1}}{\Gamma(k)\Gamma(m)}$	uniform order statistics
$\mathcal{C}(\tilde{\mu},\tilde{\sigma})$	$\frac{\tilde{\sigma}}{\pi\cdot[\tilde{\sigma}^2+(x-\tilde{\mu})^2]}$	laser beam behind screen

their mean, variance, MGF.

Bivariate Normal $(\mu_x,\mu_y,\sigma_x,\sigma_y,\rho)$

Standardized pdf

$$f(z_1,z_2)=\frac{\exp\left(-\frac{z_1^2+z_2^2-2\rho\cdot z_1z_2}{2(1-\rho^2)}\right)}{2\pi\sqrt{1-\rho^2}}$$

MGF

$$\exp\left(\frac{\sigma_1^2t_1^2+\sigma_1^2t_2^2+2\rho\cdot\sigma_1\sigma_2t_1t_2}{2}+\mu_xt_1+\mu_yt_2\right)$$

How does it 'generate moments' (differentiate, then substitute $t_1 = t_2 = 0$).

Marginal distribution of X

$$\mathcal{N}(\mu_x, \sigma_x)$$

Conditional distribution of X given $Y = \mathbf{y}$

$$\mathcal{N}\left(\mu_x + \rho \cdot \sigma_x \cdot \frac{\mathbf{y} - \mu_y}{\sigma_y}, \sigma_x \sqrt{1 - \rho^2}\right)$$

Any other **general** (continuous) bivariate distribution: Be able to find *marginal* pdf

$$f_X(x) = \int_{\text{All } y|x} f(x, y) dy$$

conditional pdf

$$f_{X|Y}(x) = \frac{f(x, \mathbf{y})}{\int_{\text{All } x|\mathbf{y}} f(x, \mathbf{y}) dx}$$

and answer any probability question about X and Y.

Sampling from univariate $\mathcal{N}(\mu, \sigma)$

$$\bar{X} \; \epsilon \; \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

$$\frac{\bar{X} - \mu}{\frac{s}{\sqrt{n}}} \; \epsilon \; \mathsf{t}_{n-1}$$

$$\frac{(n-1)s^2}{\sigma^2} \; \epsilon \; \chi^2_{n-1} \quad \text{same as} \quad \mathsf{gamma}\left(\frac{n-1}{2}, 2\right)$$

$$\frac{s_1^2}{s_2^2} \; \epsilon \; \mathsf{F}(n_1 - 1, n_2 - 1) \quad \text{(independent)}$$

Sampling from *any* other distribution (CLT):

$$\bar{X} \quad \text{is approximately} \quad \mathcal{N}\left(\mu, \frac{\sigma}{\sqrt{n}}\right)$$

Transforming $Y = g(X)$

$$F_Y(y) = \Pr\left(g(X) \le y\right)$$

or

$$f_Y(y) = f_X(g^{-1}(y)) \cdot \left|\frac{g^{-1}(y)}{dy}\right|$$

More difficult: $U = h(X, Y)$

$$F_U(u) = \Pr\left(h(X, Y) \le u\right)$$

Special case: find pdf of $X_1 + X_2$ (independent) by *convolution*.
Some transformations worth remembering (symbolically)

$\mathcal{N}(0,1)^2$		χ^2_1
$\frac{\mathcal{N}(0,1)}{\sqrt{\frac{\chi^2_m}{m}}}$	(independent)	t_m
$\frac{\frac{\chi^2_m}{m}}{\frac{\chi^2_k}{k}}$	(independent)	$\mathsf{F}(m,k)$
$-\beta \cdot \ln(\mathcal{U}(0,1))$		$\mathcal{E}(\beta)$
$\sum_{i=1}^{k} \mathcal{E}(\beta)$	(independent)	$\mathsf{gamma}(k,\beta)$
$\frac{\mathsf{gamma}(k,\beta)}{\mathsf{same}+\mathsf{gamma}(m,\beta)}$	(independent)	$\mathsf{beta}(k,m)$

Order statistics

$$f_{(i)}(x) = \frac{n!}{(i-1)!(n-i)!} F(x)^{i-1}(1-F(x))^{n-i} f(x)$$
$$f_{(i,j)}(x,y) = \frac{n!}{(i-1)!(j-i-1)!(n-j)!} \times$$
$$F(x)^{i-1}(F(y)-F(x))^{j-i-1}(1-F(y))^{n-j} f(x) f(y)$$

and $\tilde{X}$ is approximately $\mathcal{N}\left(\tilde{\mu}, \frac{1}{2f(\tilde{\mu})\cdot\sqrt{n}}\right)$ - we can deal with it exactly.

Parameter estimation

Find ML and MM estimators of one or two parameters and their variances (in case of MLE, differentiate between regular - use RCV - or not).

Find (jointly) sufficient statistic(s); these (as well as any estimators) must be functions of the X_is *only* - not containing θ or λ!