hw-lect4-1:
Consider the Bernoulli dist. with parameter pi:

$$p(x) = \pi^{x}(1-\pi)^{1-x}, \quad x=0,1 \qquad 0<\pi<1$$

a) Show that it's a distribution (prob. mass function).
b) Find the prob that x=1.

a) $p(x) \geq 0$ for $x=0,1.$ ✓

$$\sum_{\text{all } x} p(x) = \sum_{x=0}^{1} p(x) = p(x=0) + p(x=1)$$
$$= \pi^{0}(1-\pi)^{1} + \pi^{1}(1-\pi)^{0} = 1-\pi+\pi = 1 \checkmark$$

b) $\text{prop}(x=1) = p(x=1) = \pi^{1}(1-\pi)^{0} = \pi.$

This gives The param π a nice interpretation. It's The prop. of times we get x=1. If x=0,1 represent Heads and tails, Then π is The prob. of getting Heads.

hw-lect4-2:
Based on data, we have observed that x is between 0 and 1/2 about 25% of the time. Which of the following is the more reasonable distribution from which our data may have come? Show work (always!)

A) $f(x) = 2x \quad 0 \leq x \leq 1$ B) $f(x) = e^{-x} \quad 0 \leq x < \infty$

A) $\int_0^{1/2} 2x\,dx = 2\frac{1}{2}x^2\Big|_0^{1/2} = \frac{1}{4} = \boxed{25\%}$ ← It's more likely That our data have come from This distr./pop.

B) $\int_0^{1/2} e^{-x}dx = -e^{-x}\Big|_0^{1/2} = 1-e^{-\frac{1}{2}} \simeq 40\%$

hw_lect4_3: What's the prob of getting 1 or 2 boys in a sample of size 10, taken from a population in which the proportion of boys is exactly 50%?

X={"The number of boys in a sample of size 10"} has all The characteristics of being a Binomial r.v. with n=10.
Since The prop. of Boys in The pop is 0.5, Then The prob of drawing a boy is 0.5, and That's The π parameter of Binomial.

$$\therefore pr(X=1) = \frac{10!}{1!\,9!}(.5)^1(1-.5)^9 = \frac{10\,(\cancel{9!})}{\cancel{9!}}\left(\frac{1}{2}\right)^{10} = 10\left(\frac{1}{2}\right)^{10}$$

$$pr(X=2) = \frac{10!}{2!\,8!}(.5)^2(1-.5)^8 = \frac{10(9)\,\cancel{8!}}{2\,(\cancel{8!})}\left(\frac{1}{2}\right)^{10} = 45\left(\frac{1}{2}\right)^{10}$$

$$pr(X=1 \text{ or } X=2) = (10+45)\left(\frac{1}{2}\right)^{10} = 55\left(\frac{1}{2}\right)^{10}$$

hw-lect 5-1   Show That

a) $\int_0^\infty \lambda e^{-\lambda x}\,dx = 1$ $y=\lambda x \to dy = \lambda\,dx$

$LHS = \int_0^\infty \lambda e^{-\lambda x}\,dx = \int_0^\infty e^{-y}\,dy = -e^{-y}\Big|_0^\infty = -(0-1) = +1.$

b) $\sum_{x=0}^\infty \frac{e^{-\lambda}\lambda^x}{x!} = 1$ [Hint: use the Taylor series expansion for $e^{+\lambda}$]

$LHS = \sum_{x=0}^\infty \frac{e^{-\lambda}\lambda^x}{x!} = e^{-\lambda}\sum_{x=0}^\infty \frac{\lambda^x}{x!} = \frac{\sum_{x=0}^\infty \frac{\lambda^x}{x!}}{e^\lambda} = 1.$

Taylor series exp. of $e^\lambda = \sum_{x=0}^\infty \frac{\lambda^x}{x!} = \left(1+\lambda+\frac{1}{2!}\lambda^2+\ldots\right)$

c) $\int_{-\infty}^\infty \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,dx = 1$ [use $\int_{-\infty}^\infty e^{-\frac{1}{2}x^2}\,dx = \sqrt{2\pi}$]

$LHS = \frac{1}{\sqrt{2\pi\sigma^2}}\int_{-\infty}^\infty e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}\,dx = \frac{1}{\sqrt{2\pi\cancel{\sigma^2}}}\int_{-\infty}^\infty e^{-\frac{1}{2}z^2}\,\cancel{\sigma}\,dz$

$z = \frac{x-\mu}{\sigma} \Rightarrow dz = \frac{1}{\sigma}\,dx$

$= \frac{1}{\sqrt{2\pi}}\underbrace{\int_{-\infty}^\infty e^{-\frac{1}{2}z^2}\,dz}_{\sqrt{2\pi}} = 1.$

$\sum_x p(x) = 1$ for binomial will be proved, later, (and in a different way) hw-lect 5-2

Let x have a Normal dist. with params μ, σ, ie. $x \sim N(\mu, \sigma)$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \qquad -\infty < x < \infty$$

a) Find The density function ($f(z)$) for $z = \frac{x-\mu}{\sigma}$.

Hint: Start with $\int_{-\infty}^{\infty} f(x)\,dx = 1$, do not do The integral, but instead, do a change of variables until you get $\int_{-\infty}^{\infty} [\cdots]\,dz = 1$. Then $[\cdots]$ is $f(z)$.

Important

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \implies \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx = 1$$

$$z = \frac{x-\mu}{\sigma} \implies dz = \frac{1}{\sigma}\,dx$$

$$\frac{1}{\sqrt{2\pi}\cancel{\sigma}} \int e^{-\frac{1}{2}z^2} \cdot \cancel{\sigma}\,dz = 1$$

$$\therefore \int \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}\,dz = 1 \implies \boxed{f(z) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}z^2}}$$

not rigorous! but OK for now.

b) In $f(z)$, in The place where you would expect to find μ and σ, what numbers do you see?

$$f(z) = \frac{1}{\sqrt{2\pi(1)}} e^{-\frac{1}{2}\left(\frac{z-0}{1}\right)^2} \implies \boxed{\mu = 0, \quad \sigma = 1.}$$

Moral 1: if $x \sim N(\mu, \sigma)$, Then $z = \frac{x-\mu}{\sigma} \sim N(0,1)$.

Moral 2: if $x \sim$ any distr., Then This method can be used to find The distr. of any function of x (e.g. $\frac{x-\mu}{\sigma}$).

hw-lect5-3 Find The prob. of $x \le 2$ if

a) $x \sim \text{Bernoulli}(\pi = \text{arbitrary})$ b) $x \sim \exp(\lambda = 3)$

c) $x \sim \text{poiss}(\lambda = 3)$ d) $x \sim \text{Binom}(n = 10, \pi = \frac{1}{4})$

e) $x \sim N(0,1)$

a) $$\sum_{x=0}^{2} p(x) = \sum_{x=0}^{1} \pi^x (1-\pi)^{1-x} = \pi^0 (1-\pi)^1 + \pi (1-\pi)^0 = 1$$

b) $$\int_0^2 3e^{-3x}\,dx = \frac{3e^{-3x}}{-3}\Big|_0^2 = -1\left(e^{-6} - 1\right) = 1 - e^{-6}$$

c) $$\sum_{x=0}^{2} \frac{e^{-3}\, 3^x}{x!} = e^{-3}\left(\frac{3^0}{0!} + \frac{3^1}{1!} + \frac{3^2}{2!}\right) = e^{-3}\left(1 + 3 + \frac{9}{2}\right) = \frac{17}{2} e^{-3}$$

d) $$\sum_{x=0}^{2} \frac{10!}{x!\,(10-x)!} \left(\frac{1}{4}\right)^x \left(\frac{3}{4}\right)^{10-x} = \frac{10!}{0!\,10!}\left(\frac{3}{4}\right)^{10} + \frac{10!}{1!\,9!}\,\frac{1}{4}\left(\frac{3}{4}\right)^9 + \frac{10!}{2!\,8!}\left(\frac{1}{4}\right)^2\left(\frac{3}{4}\right)^8$$

$$= 1\left(\frac{3}{4}\right)^{10} + 10\,\frac{3^9}{4^{10}} + 45\,\frac{3^8}{4^{10}}$$

$$= \frac{3^8}{4^{10}}\left(3^2 + 10(3) + 45\right) = \frac{3^8 (86)}{4^{10}}$$

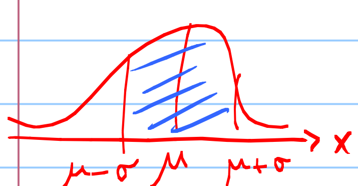
Any of These answers is acceptable.

e) $$\int_{-\infty}^{2} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}\,dx = 0.9772$$

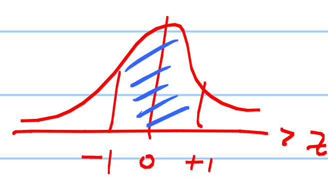
Table I.

---

hw-lect6-1 If x follows $N(\mu, \sigma)$, what's The prob. of x being within $1\,\sigma$ of μ?



=



$= .8413 - .1587 = 0.6826$

$$z = \frac{(\mu - \sigma) - \mu}{\sigma} = -1$$

hw-lect 6-2 Standardization is important in finding probs. Although it almost always refers to the change of variable $z=\frac{x-\mu}{\sigma}$, taking $N(\mu,\sigma)$ to $N(0,1)$, sometimes a different change of variable is required to obtain something that has $N(0,1)$ distr. Find the $pr(x<2)$ if $\frac{1}{x}$ has a std. Normal dist.

$$? = pr(x<2) \underset{\uparrow}{=} pr\left(\frac{1}{x} > \frac{1}{2}\right) = pr(z>0.5) = 1 - pr(z<0.5) = 1 - \cancel{.9772}\ 0.6915$$

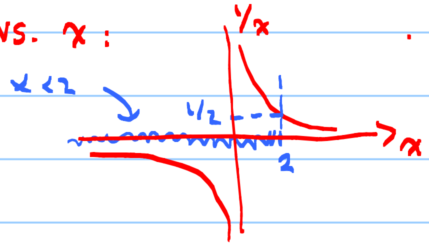
Standardize

$$\cancel{= .0228}$$

$$0.3085$$

I've asked the grader to give full credit for this, But

Technically, this soln is incomplete/wrong, because $x<2$ is equivalent to $\frac{1}{x} > \frac{1}{2}$ or $\frac{1}{x} < 0$. You can see this by looking at the graph of $\frac{1}{x}$ vs. x:



So $pr(x<2) = pr\left(\frac{1}{x} > \frac{1}{2}\right) + pr\left(\frac{1}{x} < 0\right)$

$= \text{above} + 0.5$

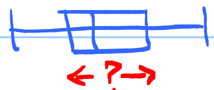
hw-lect 6-3 Another of the named distributions is the so-called power-law dist. It's formula is $f(x) = \alpha x^{\alpha-1}$, $0<x<1$, where $\alpha>0$ is its parameter.

a) Find the n^{th} percentile of that distribution. Hint: The answer will depend on α and n.

$$\int_0^{n^{th}\text{ percentile}} \alpha x^{\alpha-1}\,dx = \frac{n}{100} \Rightarrow \cancel{\alpha}\cdot \frac{x^{\alpha}}{\cancel{\alpha}}\Bigg|_0^{n^{th}\text{ percentile}} = \frac{n}{100} \Rightarrow (n^{th}\text{ percentile})^{\alpha} = \frac{n}{100}$$

$$n^{th}\text{ percentile} = \left(\frac{n}{100}\right)^{1/\alpha}$$

b) How long is the box portion of the corresponding boxplot? Hint: The answer will depend on α.



$? = 75^{th}\text{ percentile} - 25^{th}\text{ percentile}$

$= \left(\frac{75}{100}\right)^{1/\alpha} - \left(\frac{25}{100}\right)^{1/\alpha} = \left(\frac{3}{4}\right)^{1/\alpha} - \left(\frac{1}{4}\right)^{1/\alpha}$

hw_lect6_4

Consider ONE of the 2 continuous random vars, and ONE of The 2 discrete (categorical variables in the data you collected. Make comparative boxplots for the continuous variable for each level of the discrete variable. E.g. if the discrete var. has 4 levels, then you need to show 4 boxplots for the cont-variable, all on the same plot, side-by-side. Interpret/discuss them. By R.

The answers will vary across students, but The code will be something like This :

```
dat = read.table( ... ).
x1 = dat [,1]      # 1st categorical/discrete var.
x2 = dat [,2]      # 2nd      "       "
x3 = dat [,3]      # 1st continuous var.
x4 = dat [,4]      # 2nd      "       "
boxplot( x3[x1 == A], x3[x1 == B], ..., x3[x1 == Z])
                   # where A, B, ..., Z are the levels of x1.
```

Interpretations include The "shape" of The boxplot (e.g., is There a skew), The width (which ones are wider, why?), and The relative position of The boxplots (is There a difference between The groups?)