↑ Remember when
of hypotheses
"beast mode"? Remember when we did a continuum we diel $,$ and $,$ u)e 1($we can be used\n $\frac{1}{\pi}$$ un Let's build this up carefully. Support
if you have a model that that the up
if you have a model that the model m
it is ressure the parameters be ave
two
might build this Remember when we did
of hypotheses, and
"beast mode"? Well,
some gamer "terms"
beast mode is G
feast mode is G and above Remember when we did a continuum let's build this up corefully. So
of hypotheses, and we called it was have a model that that
"beast mode"? Well, I hoked up
some gamer "terms and above the parameters" in it. The model
bea pressure some gamer terms, and above
beast mode is God mode.
To do is worthy of divinity,
another step up. I am على مراجع
نائم كم نم Some gomer terms and above Atthough
 τ don't think what we are about

To do is worthy of divinity, it is

another step up. I am deping

to call it I descient has two independent hypotheses hypotheses : of *hypotheses*, and we called it
"beast) mode"? Well, I looked up
some gome "terms and above Athough $p = a t + b$
beast) mode is God mode. Athough $p = a t + b$
fo do is worthy of divinity, it is
nother step up. I am (going or su $-\frac{1}{2}$ A Muitti-Uimensiona, Continuum
of Hypotheses
And when we are done deriving paper was written paper was written and when we are done derived that independent hypotheses:

A Multi-Dimensional Continuum

And when we are done deriving

Bayes Theorem in this case in a proper was written paper was written

Bayes Theorem in this case in coult go even for ther and apply
Havesian conjugates that
distributions with two parameters,
such as ul and o that characterize The other independent by pothes's Also, what I am about to do will Sayesian conjugates 48 will authoredin ¹⁷⁸⁷ authoredin ¹⁷⁸⁸ Also, what I am about to do will promised a course that didn't
have cale as a prerego we
introduced the cale we needed as possibilities! We The other independent
apply
apply
eferts
acterize The other independent
as a suit of there are
d as possibilities!
Institutes! So there are 4 motially exclusive (have calc as a prerez. We
introduced the calc we needed as possibility
it came up. But this new subject $||$ the possibility $\overline{\mathcal{J}}_{\mathcal{J}}$ Hes!
E H, NE
E WH, NE it came up. But this new subject $||$ H,
takes us into multi-variance calculus... $||$ H, takes us into multi-variable calculus!!

↑ We are going to need a new version of $\left| \frac{7\#EN}{9F} \right|$ we EQUATE THOSE TWO WAYS $P(H_i, E_j)$ data) $P(data)$ To make things more general, I
am going to make Hand at
part of a list of NH hypotheses, H_0 , H_1 , H_2 , ..., H_{ν} -z, $H_{\nu-1}$ am going to make H and γ = P (data/H; E;) P(H; E;)
part of a list of N_H hypotheses, THIS IS ALL OLD HAT. WE SOLVE:
Ho, H, Hz, ..., H_{γ} -z, H_{γ} -l, H_{γ} -z, H_{γ} -l, H_{γ} -z, H_{γ} -l, H_{γ} -z, H_{γ})
[] part of a list of N_{m-2} , N_{m-1}
and ϵ part of a list of N_{ϵ} P(H_{i)} ϵ , ϵ data) = P(data) + i, ϵ bypotheses, $\epsilon_0, \epsilon_1, \epsilon_2, \cdots, \epsilon_{N_{\epsilon}-2}$, $\epsilon_{N_{\epsilon}-1}$ or a list of N_{m-2} , N_{n-1}
 \neq p_{a+1} , H_{a+2} , H_{n-1}
 \neq p_{a+2} , \neq p_{a+1}
 \neq p_{a+2} , \neq $\$ AND GET $P(E_i, E_j \mid data)$ $P(data)$
 $P(E_i, E_j \mid data)$ $P(data)$
 $P(At_i, E_j \mid data)$ $P(data)$
 HIS is ALL OLD HAT, WE SOLI
 $(H_i, E_j \mid data) = \frac{P(data/H_i, E_j) P(iH_i, E_j)}{P(data)}$
 $FINALLY$ WE EXPAND THE DENOMINA

AND GET
 $P(H_i, E_j \mid data) = \frac{P(data|H_i, E_j) P(H_i, E_j)}{P(H_i, E_j) P(H_i, E_j)}$ Boyes theorem in this case OF COURSE, WE ARE SOMETIMES Ca / 11, J C J V (111)

OLD HAT. WE SOLU:
 $\begin{aligned} \mathcal{P}(data/ H_i, E) P(H_i) \ = \mathcal{P}(data/ H_i, E) P(H_i) \ = \mathcal{P}(data/ H_i, E) P(H_i, E) \ = \mathcal{P}(data/ H_i, E) P(H_i, E) \ = \mathcal{P}(data/ H_i, E) P(H_i, E) \ \mathcal{W} \in ARE \quad SOMET/MES \ \mathcal{W} \neq \mathcal{W} \end{aligned}$ comes from writing ,,,,, BLOPPY (AS YOU HAVENOW SEEN IN $P(H_i, E_j \cap data)$ MULTIPLE CONTEXTS], AND WE USE i AND] IN THE DENOMINATOR, EVEN $= P(H_i, E_i / data) P(data)$ THOUGHTHAT i AND THAS NOTHING
OTE A ZND WAY TO DO WITH THE " I AND)
IN THE NUMERATOR. THEN WE $P(H_i, E_j)$ $\begin{pmatrix} 2 & 4 \\ 6 & 1 \end{pmatrix}$ $= P(data/H_i, E_i) P(H_i; E_i)$ To, H_1 , H_2 , $...$, H_{n-2} , H_{n+1}
 $d \in part$ of a list of N_e
 H_{n+2} , H_{n+1}
 $f(H_1) \in \text{Joda}(H_2) \cup \text{Jida}(H_1)$
 $= \text{Joda}(H_2)$
 $= \text{Jaba}(H_2)$
 $= \text{Jaba}(H_1)$
 $= \text{Jaba}(H_1)$
 $= \text{Jaba}(H_1)$
 $= \text{Jaba}(H_1)$
 $= \text{Jaba}(H$ $HAVE$
P(H_i) E_j / data) = $\frac{P(dada/H_i, E_j)P(H_i, E_j)}{P(dada/H_i, E_j)P(H_i, E_j)}$

↑ we are not in "God-mode" yet. We just In the numerator μ_{ij} σ_{ij} , $\Delta \mu_{ij}$ and $\Delta \sigma_j$ Leve are not in "God-mode" yet. We just
have a fancy way of covering Ny. Ne discrete,
mutually-exclusive pairs of hypotheres: a fancy way of covering Ny. NE
ly-exclusive pairs of hypotheres: can be whatever you $\frac{\partial}{\partial k}$ denominator, because the sums have to range over $P(H_i, E_j|data) = \frac{P(data|H_i, E_j)P(H_i, E_j)}{\sum_{i=0}^{N_{i-1}}\sum_{j=0}^{N_{i-1}}P(data|H_i, E_j)P(H_i, E_j)}$ In the numerator, μ_{ij} σ_j
can be whatever you like. In
denominator, because the sems have
all possible values of μ and σ_j , we must have ay of covering W_y . No discrete, can be whatever you like. In the
pairs of "hypotheres: denominator because the sems have to range over
= $\frac{P(data|H, E, P(H, E,))}{P(data|H, E, P(H, E,))}$
 $\sum_{i=0}^{N_y - 1} \sum_{j=0}^{N_z - 1} P(data|H, E, P(H, E,))$
 $P(H_i, E_j | data) = \frac{P(H_i, E_j | data | H_i, E_j) P(H_i, E_j)}{P(A \cdot E_j | data | H_i, E_j) P(H_j, E_j)}$
The Multi-Dimensional Continuum of course, if μ and σ are σ $\frac{1}{\frac{N}{M}}$, if μ and $\frac{N}{M}$ are
introduced or are continuous, the only way this is going to be a precise expression is if Imagine that Hi represents the chance between μ_i and μ_i + $\Delta \mu_i$ and μ_i this by writing the denominator as E_j represents the chance that a continuous variable O is between $I_{\text{min}}^{\text{max}}$
 $I_{\text{min}}^{\text{max}}$
 N_{min} σ_{j} and σ_{j} + 10, σ_{j} , σ_{j} rge. We express
The denominator as
ata / Uk, G)
P / Uk, G) SUS chance that a

only between

is between $N_H \rightarrow N$
 $N_H \rightarrow N$
 There is a special symbol from
multivariable calculus that capture Then multivariablethe E_j represents the
continuous variable
of and $5 + 4$
Then $P(U_i) = \int d \alpha t \alpha$
 $= P(d \alpha t \alpha / \mu)$
 $= \sum_{k=0}^{N} \alpha^{k-1} \sum_{k=0}^{N-1} \alpha^{k-1}$ $P(\mu_i)$ $T/data$ $\triangle\mu\triangle\sigma$, V^{olum}
 $P(data/\mu_i)$ $T/data/\mu_i$ $T/data/\sigma$, $P(\mu_i)$
 Tae interpretation is that σ \sim μ_{min} W $\frac{P(data|u; \sigma) P(y| \sigma) dy d\sigma}{P(data|u; \sigma) P(y| \sigma) dy d\sigma}$
 $\frac{P(data|u; \sigma) P(y| \sigma) dy d\sigma}{P(data|u|, \sigma) P(y| \sigma) dy d\sigma}$
 $\frac{P(data|u; \sigma) P(y| \sigma) dy d\sigma}{P(data|u|, \sigma) P(y| \sigma) P(y| \sigma) dy d\sigma}$ $\begin{bmatrix} \frac{1}{2} & \frac{1}{2$ $P|data(y|<)P(y|<)$

↑ Often, the lower and upper limits of Often, the lower and upper limits of Plu, σ /data) $\Delta\mu\Delta\sigma$ int integration μ min, μ may σ min, and σ the denominator as we write $=$ P(data/u $=$ P/u S)W16 $\mathcal{P}(data|y,0)P(y,0)dyd0$ $\int\int P(d\tau|{\rho}^{\prime})\mu(\sigma)P(\mu,\sigma)$ dudo Also, in the numerator, Compare this with Donovan & Mickey Hiso, in the numerator, we have the $Eq.$ 12.10 on $Eq.$ 12.10 on $p.$ 180. It is the exact same except that
they don't have the sul and so can be anything, not some discrete $\frac{1}{2}$ is the exact same excepted the $\frac{1}{2}$
set of values that becomes larger they don't have the $\frac{1}{2}$
and larger as $N_H \rightarrow \infty$ and $N_E \rightarrow \infty$ $m_V / \frac{1}{2} / \frac{1}{2}$ the numerato and larger as N_H $\rightarrow \infty$ and N_E $\rightarrow \infty$ Then instead of $\frac{y}{\sqrt{p}}$ You can of course cancel thase you can of course can rccp1 In
1 and 2
6 Have
6 Have $P(u_i) = \int d \alpha t \alpha \int \Delta \mu_i \Delta \sigma_j$
= $P(d \alpha t \alpha / \mu_i, \sigma) P(\mu_i \sigma) \Delta \mu_i \Delta \sigma_j$ and in the interpretation $\sum_{k=0}^{N_d-1}\sum_{l=0}^{N_d-1}P(d_n t a / \mu_k, \epsilon_l)$ of the probability density. we have He aid in the interpretation
of the probability density.
Now I will cancel appear offsince.. Now I will cancel an ao off since
it is in the numerator on both sides, and we know how to put it in the numerator on both sides and we know how to put it
back if we want to resture the density.

↑ Bayesian Conjugates The version of Bayes' Theorem at left with Gaussian Data becomes $P(\mu)$ f Bayes
 $\int c \int d \tau d \tau$ On the last four pages, we derived =datau Plu,G(data) SSP(data/e,T)PM) dude =datau Let us remember the , $=\frac{\mathcal{P}(data/\mu_{0}|\sigma)\mathcal{P}(\mu_{0}|\sigma)}{\int\mathcal{P}(data/\mu_{0}|\sigma)\mathcal{P}(\mu_{0}|\sigma)dud\sigma}$ of chapters 10-12. On the last four pages, we derived
 $P(\mu_{1} \sigma / \text{data}) = \frac{P(\text{data}/\mu_{1} \tau)}{\int (P(\text{data}/\mu_{1} \sigma / \text{p})/\mu_{1} \sigma)} = \frac{P(\text{data}/\mu_{1} \tau)}{\int (P(\text{data}/\mu_{1} \sigma)P(\mu_{1} \sigma) \text{d}\mu \sigma} = \frac{P(\text{data}/\mu_{1} \tau)}{\int (P(\text{data}/\mu_{1} \sigma)P(\mu_{1} \sigma) \text{d}\mu \sigma} = \frac{P(\text{data}/\mu_{1} \$ We did this in the midst of
Chapter 12 because our likelihoods
in Chapter 12 involve two
garameters, Our canonical example agical pri
Vesian coni $\overline{\mathcal{E}}$ $\frac{1}{\sqrt{2}}$ involve two called SaveSian conjugates whose magical property is that is bacteria survival times: the functional form of the $\mathcal{P}(\chi/\mu)$ Let G
 $\begin{array}{rcl}\n\mathbf{A} & \mathbf{Gaussian} & \mathbf{Dada} & \mathbf{faussian} & \mathbf{f} & \mathbf{f} & \mathbf{f} \\
\mathbf{f} & \mathbf{f} & \mathbf{f} & \mathbf{f} \\
\mathbf{f} & \mathbf{f} & \mathbf{f} & \mathbf{f} \\
\mathbf{f} & \$ $\begin{array}{l} \mathcal{A}(M,\sigma)=\frac{1}{12\pi^2}\frac{1}{\sigma^2}\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{1}{2}\left(1-\mu\right)^2\right)\left(1-\frac{$ $\frac{18}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\int_{0}^{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac{1}{\sqrt{2\pi}}\frac$ $\frac{1}{\frac{1}{\sqrt{m}}\sum_{i=1}^{m}x_i}$ (but with different parameters) Tt turns out that the algebra (out with different paranger It turns out that the algebra as the functional form of
is a little tidier if we let $P(\chi/\mu,\sigma) = \frac{1}{\sqrt{2\pi}} \frac{C}{\sigma}$
survival mean
fine
 χ^2 furns out that
is a little tidier if
 $\alpha = \frac{1}{\sqrt{2\pi}}$ and write the $z=\frac{1}{s^2}$ and write the likelihood as
 $P(\chi/\mu, z) = \sqrt{\frac{z^2}{z\bar{p}'}}e^{-\frac{1}{z^2}(\chi-\mu)^2/z}$ as the functional fam of
the prior. So that is
what we are seeking for product of the likelihed and
the prior) is the same
(but with different parampters)
as the functional form of
the prior. So that is
what we are seeking for
the likelihood at

And now I $\begin{aligned} \n\begin{aligned}\n\text{and} & \mathcal{F}(\text{ref}) \\
\text{and} & \mathcal$ j ust have to write the likelihood for n new data points $\not\! k$, is mag ical prior down, and there caussian and independently and Gaussian and independently and
identically distributed (i.i.d.) and
the prior is "prior P(M,) ⁼ do the nasty algebra to show identically distri
that it is conjugate to the Gaussian likelihood. Without further ado: saussian likelitoot, $w + hovf$ further are:
 $w + v^2 = P(u, v) = P(u, v)P(v)$
 $w + v^2 = P(u, v) = P(u, v)P(v)$
 $w + v^2 = 0$
 $P(u, v) = \sqrt{\frac{n_0 v}{\pi}} \approx \frac{-n_0 v}{\pi} (u - u) \frac{v^2}{\pi}$
 $P(u, v) = \sqrt{\frac{n_0 v}{\pi}} \approx \frac{-n_0 v}{\pi} (u - u) \frac{v^2}{\pi}$
 $w = 0$
 $w = 0$
 w $P(\mathcal{P}) = P(\mathcal{P}) = P(\mathcal{P}(\mathcal{P})) P(\mathcal{P})$ and $P(z)$ $\tau | x \sim Ga\left(\alpha + \frac{n}{2}\right), \quad \beta + \frac{1}{2}\sum_{i}(x_i - \bar{x})^2 + \frac{nn_0}{2(n+n_0)}(\bar{x} - \mu_0)^2\right).$ for P(2) is the Infordan's notation the curly N represents This formula This formula for $P(6)$ is the π fordon's notation the curil N represent
Same one that showed up in a "normal" distribution, and is Juhat we call a Same one that snowed of $\begin{array}{l|l} \hline \end{array}$ Gaussian. The "Ga" represents a gamma
Chapter 11 for the prior that distribution and is identical to the gamma distribution Chapter 11 for the prior that students and is identical to the grammetiches

is conjugate to Poisson distributions and that was also in my Bayesian course

was introduced. and if $\frac{1}{2}$ do not trust these formulas (Jor Chapter 11 for the prior that distribution and is identical to the gamma district is conjugate to Poisson distributions and that was also in my Bayesian conjugate to Poisson distributions and that was also in my Bayesian Donovan and Mickey have in Eq 11. ⁸ on p is conjugate to Poisson distributions and that was also in my Bayesian conjugates
In the Conjugate to Poisson distributions and that was also in my Bayesian conjugates rad students write the course notes he posts);
let a lone claim to understand them,
would need to do the derivation myself. ↑ $\begin{array}{ccc} \text{I} & \text{$ $f^{\text{2me}}_{\text{1em}}$ an count on in
not including -
- S But it you
challenge, y to be frank, I havent done it with what you can count on in this handout is everything
yot myself. In Michael I Jardany 40 but not including this last page nk
!se/F. what you can count on in this handout is everything $\begin{array}{c|c} \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \overline{0} & \overline{0} & \overline{0} & \overline{0} \\ \end{array}$ ²⁰¹⁰, on Stat 260 handout for February 8
p.6 he says that if the But it you want a serious
challenge, you can verify the rest (c)